

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G09G		A2	(11) International Publication Number: WO 99/35633
		(43) International Publication Date: 15 July 1999 (15.07.99)	
(21) International Application Number: PCT/US99/00086		Charles, M. [US/US]; 667 N. Bradford Street, North Andover, MA 01845 (US).	
(22) International Filing Date: 4 January 1999 (04.01.99)		(74) Agents: VOCK, Curtis, A. et al.; Duft, Graziano & Forest, P.C., Suite 140, 1790-30th Street, Boulder, CO 80301-1018 (US).	
(30) Priority Data: 60/070,512 6 January 1998 (06.01.98) US 60/100,046 11 September 1998 (11.09.98) US		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications US 60/070,512 (CIP) Filed on 6 January 1998 (06.01.98) US 60/100,046 (CIP) Filed on 11 September 1998 (11.09.98)			
(71) Applicant (for all designated States except US): THE VIDEO MOUSE GROUP [US/US]; 18 Whitney Avenue, Westwood, MA 02090-2948 (US).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): FREY, Robert, D. [US/US]; 19 Still River Road, Bolton, MA 01740 (US). GREALISH, Kevin [US/US]; 18 Whitney Avenue, Westwood, MA 02090 (US). VOCK, Curtis, A. [US/US]; 1279 Elder Avenue, Boulder, CO 80304 (US). MARSHALL,			

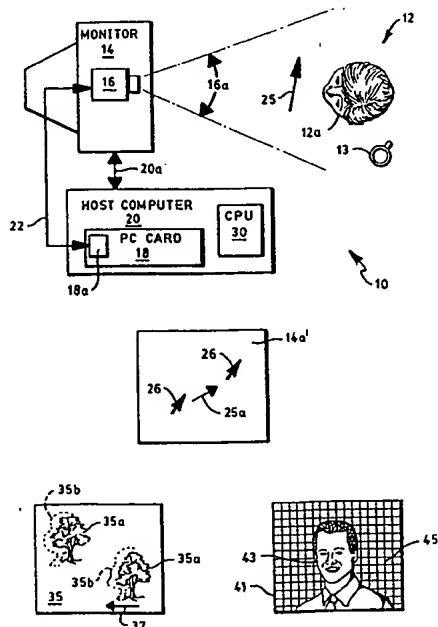
Published

Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN MOTION FOLLOWING COMPUTER MOUSE AND GAME CONTROLLER

(57) Abstract

A human motion following controller is provided by the invention to augment motion of items (e.g., computer cursor or scene view) shown on a computer display. The display is coupled to the computer which controls positioning of the items through operating system controls. A camera captures frames of data corresponding to a first image of at least part of a user (e.g., eyes, hands) at the computer display. Signal processing electronics coupled to the camera (a) detects differences between successive frames of data corresponding to motion of the first image, and (b) communicates differences information to the computer to reposition display of the items through the operating system controls. The items are thus repositioned on the display by an amount corresponding to the motion of first image.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HUMAN MOTION FOLLOWING COMPUTER MOUSE AND GAME CONTROLLER

5 A portion of the disclosure of this patent document contains material which is
subject to copyright protection. The copyright owner has no objection to the
facsimile reproduction of any one of the patent document or the patent disclosure, as
it appears in the Patent and Trademark Office patent file or records, but otherwise
reserves all copyright rights whatsoever.

Related Applications

10 This application is a continuing application of commonly-owned and co-
pending U.S. provisional application number 60/070,512, filed on January 6, 1998,
and U.S. provisional application number 60/100,046, filed on September 11, 1998,
each of which is incorporated herein by reference.

Background

15 The primary human interfaces to today's computer are the keyboard, to enter
textual information, and the mouse, to provide control over graphical information.
20 These interfaces help users with word processing, presentation software, computer
aided design packages, spreadsheet analyses, and other applications. These
interfaces are also widely used for computer gaming entertainment; though they are
often augmented or replaced by a joystick.

25 In daily use of business software applications, control of cursor position on
the screen requires that the user remove his/her hand from the keyboard in order to
use the standard mechanical mouse. The use of the mouse introduces several issues.
In a desk environment, the mouse requires maintenance of space on the desk area.
The mouse cord must also remain free from obstruction to facilitate movement.
30 Additionally, the use of the mouse is a major contributing factor of carpal-tunnel

syndrome. It would be advantageous therefore to find an alternative to the mechanical mouse.

5 In computer gaming, game complexity generally requires control of the (i) mouse and keyboard, or (ii) joystick and keyboard. Further, gaming applications usually require control in several axes of motion, including forward motion, reverse motion, left turn, right turn, left strafe (slide), right strafe, upward motion, downward motion. To further complicate game maneuvers and control, many games permit viewing (within the game environment) in directions different from
10 that in which the vehicle (e.g., the car, or person, simulated within the game) is moving, including up, down, left and right. These many complexities of motion in fact increase or modify the complexity and enjoyment of the game.

Nevertheless, these complexities require that the user have utmost dexterity
15 and control of his/her body. One object of the invention, therefore, is to offer alternative approaches to human-computer interfaces for those incapable of using standard devices (e.g., mouse, keyboard and joystick) such as due to disability.

Another object of the invention is to provide an alternative input device for
20 laptop computers. Laptop computers are used in locations which do not allow the use of a mouse, in airplanes or during business meetings in which there is no room to operate the mouse. Through the use of either a clip on camera or a camera built into the laptop display, the laptop user can control the mouse position or use the camera for teleconferencing while on the road.

25

Other objects of the invention are to replace or augment existing human computer interfaces to facilitate enhanced gaming and/or control within game environments.

30 In the prior art, certain systems exist which attempt to reduce the amount of physical interaction required with game controllers. However, such systems are

prohibitively expensive to the general public as their costs are driven by techniques and algorithms which detect user head motion based upon a detectable target worn by the user. Other costly and cumbersome systems require the user to wear apparatus which emits or detects a signal. It is thus one other object of this invention
5 to provide a system which detects user motion without the aid or augmentation of artificial devices placed on the user operator.

Another object of the invention is to provide a means of human control of a graphical computer interface through the physical motion of the user in order to
10 control the activity of a cursor in the manner usually accomplished with a computer mouse.

A further object of the invention is to provide additional degrees of freedom in the human computer interface in support of computer games and entertainment
15 software.

Yet another object of the invention is to provide dual use of teleconferencing and video electronics with gaming and computer control systems.

20 These and other objects will be apparent in the description which follows.

Summary of Invention

As used herein, "cursor" means a computer cursor associated with a
25 computer screen. "Scene view" means the view presented on a computer display to a user. For example, one scene view corresponds to the scene presented to a user during a computer game at any given moment in time. The game might include displaying a scene whereby the user appears to be walking in a forest, and through trees. In another example, a cursor might also be visible in the scene view as a
30 mechanism for the user to select certain events or items on the scene (e.g., to open a door in a game, or to open a folder to access computer files).

As used herein, "camera" refers to a solid state instrument used in imaging. Typically, the camera also includes optical elements which refract light to form an image on the camera's detector elements (typically CCD or CMOS). For example,
5 one camera of the invention derives from a video-conferencing camera used in conjunction with Internet communication.

In one aspect, the invention provides systems and methods to control computer cursor position (or, for example, the scene view or game position as
10 displayed on the computer display) by motion of the user at the computer. A camera rests on or near to the computer, or built into the computer, and connects therewith to collect "frames" of data corresponding to images of the user. These images provide information about user motion, over time. Software within the computer assesses these frames and algorithmically adjusts cursor motion (or scene
15 view, or mouse button, or some other operation of the computer) based upon this motion. The motion may be imparted by up-down or left-right motion of the user's head, by the user's hands, or by other motions presented to the video camera (such as discussed herein). In one aspect, a close up view of the users facial features is used to impart a translation in the cursor (or scene view) even through the features
20 in fact rotate with the user's head. In yet another aspect, the rotation is used to generate a corresponding rotation in computer game scene imagery.

In one aspect, the invention also provides a human factors approach to cursor movement in which the user's rate of motion determines the relative motion of the
25 cursor (or scene view). By way of example, the faster the user's head travels over a set distance, the further the corresponding cursor movement over the same time period.

In other aspects of the invention, the camera is either (a) a visible light camera
30 utilizing ambient lighting conditions or (b) a camera sensitive in another band such as the near infrared ("IR"), the IR, or the ultraviolet ("UV") spectrum. In the latter

case (b), the illumination preferably emanates from a source such as an IR lamp which is beyond human sensory perception. The sensor is typically mounted facing the user so as to capture a picture of the user's face in the associated electromagnetic spectrum. The lamp is typically integrated with the camera housing so as to facilitate
5 production and ease of consumer set-up.

In one aspect, a system of the invention provides an IR camera (i.e., a camera which images infrared radiation) to image the user's face and to gauge the user's stress level associated with a game on the computer. As the user's intensity increases
10 (such as in a fast moving computer game using a joystick or the methods discussed herein), the system detects increased heat intensity on the user's face, forehead or other body part by the imagery of the IR camera. This information is fed back into the game processor to provide further enhancement to the game. In this manner, the system gauges the user's reaction to the game and modifies game speed or operation
15 in a meaningful way. For example, suppose such a system determined that a particular user was bored of the present game speed (a determination of boring can be made by assessing low IR output over large portions of the user's face). The computer processor and game software can then cooperate to increase the gaming speed and thereby increase this particular user's stress. Games of the invention are
20 thus made and sold to users with varying intelligence, age and/or computer familiarity; and yet the system always "pushes the envelope" for any given user so as to make the game as interesting as possible, automatically.

In accord with one aspect of the invention, images captured by the sensor are
25 processed by a digital signal processor ("DSP") located either (a) in a PC card within the host computer or (b) in a housing integrated with the sensor. In case (a), sensor frames are sent to the PC card; and detected user motion (sometimes denoted herein as "difference information") is communicated to the user's operating system via a PCI (or USB or later standard) bus interface. These difference information
30 commands are interpreted by a low overhead program resident at the user's main processor, which either updates the cursor position on the screen or provides motion

information to the user's computer game (e.g., so as to change the scene view). In case (b), the DSP is contained within the camera housing; and frames are processed local to the camera to determine difference information. This information is then transmitted to the computer by a cable that connects to a bus port of the computer so that the host processor can make appropriate movements of the cursor or scene view. In another aspect, the DSP is mounted in the camera housing such that the camera/signal processing subsystem produces signals which emulate the mouse via the mouse input connector.

In an alternative configuration, frames of image data are sent directly to the host computer through the computer bus; and that image data is manipulated by the computer processor directly. With increasing computer processing speed, it is expected that sensor data frames can be sent directly to the host processor for all processing needs, in which case the PC card and/or separate DSP are not required. Although this is possible today, the update rates are likely too slow for practicality. Once GHz processors are on the market, a separate DSP may no longer be needed.

In one aspect of the invention, pixel format or pixel density of the camera drives the accuracy of the system. Higher pixel density in the image of the user's face, for example, increases the attainable resolution and cursor control (or the attainable control of scene view motion). Camera formats of 240 vertical by 320 horizontal generally provide satisfactory performance. The number of pixels that may be utilized is determined by system cost factors. Greater numbers of pixels require more powerful DSPs (and thus more costly DSPs) in order to process the image sequences in real time. Current technology limits the processing density to a 64x64 window for consumer electronics. As prices are reduced, and power increases, the densities can increase to 128x128, 256x256 and so on. While 64x64 density is satisfactory for general household users, a higher fidelity system using a greater number of pixels is possible, in accord with the invention, for higher end applications at a proportionally higher cost. Non-square pixel formats are also possible in accord with the invention, including a 64x128 detector array size.

In one aspect, the data transfer rate from the camera is 30 frames/second at 240x320 pixels per frame. Assuming eight bits per pixel, the digital data transfer rate is therefore 18.432 megabits/second. This is a fairly high transfer rate for consumer products using current technology. While the data transfer can be either analog or digital, the preferred method of image data transfer for this aspect is via a standard RS170 analog video interface.

In accord with one aspect, a system of the invention defines two imaging zones (either within a single camera CCD or within multiple CCD cameras housed within a single housing). One imaging zone covers the user's head; and the other covers the user's eyes. This aspect includes processing means to process both zones whereby movement of the user's head provides one mechanism to control cursor movement (or scene view motion), and whereby the user's eyes provide another mechanism to control the movement. In essence, this aspect increases the degrees of freedom in the control decision making of the system. By way of example, a user might look left or right within a game without moving his head; but by assessing movement of the user's eyes (or the pupils of those eyes), the scene view can be made to rotate or translate in the manner desired by the user. Further, a user might move his head for other reasons, and yet not move her eyes from a generally forward looking position; and this aspect can assess both movements (head and eyes) to select the most appropriate movement of the cursor or scene view, if any.

In another aspect, a system of the invention utilizes a camera with zoom optics to define the user's pupil and to make cursor or scene views move according to the pupil. In another aspect, the system incorporates a neural net to "learn" about a user's eye movements so that more accurate movements are made, over time, in response to the user's eye movement.

In still another aspect, a neural net is used to learn about other movements of the user to better specify cursor or scene view movement over time.

In yet another aspect of the invention, a system is provided with two CCD arrays (either within a single camera body or within two cameras). The arrays connect with the user's computer by the techniques discussed herein. One CCD array is used to image the user's head; and the other is used to image the user's body. Motion of the user is then evaluated for both head and body movement; and cursor or scene view movement is adjusted based upon both inputs.

In another aspect of the invention, a single CCD is used to image the user. However, alternate frames are zoomed, electronically, so that one frame views the user's head, and the next frame views the user's eyes. With the algorithm discussed herein, these separate frame sequences (one for the eyes, one for the head) are processed separately and evaluated, together, to make the most appropriate cursor or scene view movement. If for example, the system clocks at 30Hz, then one set of frame sequences operates at 15Hz, and the other at 15Hz. However, the advantage is that two movement information sets can be evaluated to invoke an appropriate movement in the cursor or scene view.

Those skilled in the art should appreciate that different frame rates can be used; and frame rates for either sequence (head or eyes) can occur at different rates too. Further, the separate frame sequences can utilize other body parts, e.g., the head and the hand, to have two movement evaluations. Alternatively, a separate camera (or CCD array) can be used to image other body parts, for example one camera for the head and one for the hand.

25

The invention also provides methods for shifting cursor or scene views in response to user movement. In one aspect, the scene view shifts left or right when the user shifts left and right. In another aspect, the scene view rotates when the user's head rotates. This last aspect can be modified so that such rotation occurs so long as the eyes do not also rotate (in this situation, the user's head rotates, indicating that she wishes the scene view to rotate; but the eyes do not, indicating

30

that she still watches the game in play). In another aspect, the scene view rotates in response to the user's hand rotation (i.e., a camera or at least a CCD array of the system is arranged to view the player's hand).

5 In another aspect, the invention provides a multi-zone player gaming system whereby the user of a particular computer game can select which zone operates to move the cursor or the scene view. By way of example, the system can include one zone corresponding to a view of the user's head, where frames of data are captured by the system by a camera. Another zone optionally corresponds to the user's hand.
10 Another zone optionally corresponds to the user's eyes. Each zone is covered by a camera, or by a CCD array coupled within the same housing, or by optical zoom zones within a single CCD, or by separate optical elements that image different portions of the CCD array. By way of example, two zones can be covered with a single CCD array (i.e., a camera) when the zones are the user's head and eyes. The
15 camera images the head, for one zone, and images the eyes in another zone, since the zones are optically aligned (or nearly so). However, two cameras (or optionally two CCD arrays with separate optics) can view two zones such as the user's head and the user's hand. Combinations of zones is also possible and envisioned in accord with the invention.

20

Zones in a single camera can also be identified by the computer by prompting the user for motion from corresponding body parts. For instance, the computer identifies the head zone by prompting the user to move his head. Then the computer identifies the foot zone by having the user move his foot. Once the
25 zones are identified, the motion of each of these individual zones is tracked by the computer and the regions of interest in the camera image related to the zones moved as the targets in the zones move with respect to the camera.

In one aspect, the invention provides a system, including a camera and edge
30 detection processing subsystem, which isolates edges of the user's body, for example, the side of the head. These edges are used to move the cursor or scene

view. For example, if the left edge of head is imaged onto column X of one frame of the CCD within the camera, and yet the edge falls in column Y in the next frame, then a corresponding movement of the cursor or scene view is commanded by the system. For example, movement of the edge from one column to the next might
5 correspond to ten screen pixels, or other magnification. In one aspect, this magnification is selected by the user. Up and down motion can also be detected by similar edge detection. For example, by imaging the user's chin, an edge movement in the up or down dimension is formed (e.g., if the bottom edge of the chin moves from one row to the next, in adjacent frames, then a corresponding movement of the
10 cursor or scene view is made - magnification again preferably set manually with a default starting magnification). Other images can also serve to define edges. For example, in one aspect, a user's eyelash can be used to move the cursor (or scene view) up or downwards; though typically the eye blink is used to reset the cursor command cycle.

15

In one aspect, an optical matched filter is used to center image zones onto the appropriate images. For example, as discussed above, one aspect preferably utilizes 64x64 pixels as the image frame from which cursor motion is determined. Many cameras have, however, many more pixels. These 64x64 arrays are therefore
20 preferably established through matched filtering. By way of example, an image of a standard pair of user's eyes is stored within memory (according to one aspect of the invention). This image field is cross-correlated with frames of data from the actual image from the camera to "center" the image at the desired point. With eyes, specifically, ideally the 64x64 sample array is centered so as to view both eyes within
25 the 64x64 array. Similarly, to process sequences of head data, a standard head image is stored within memory, according to one aspect, and correlated with the actual image to center the head view.

Those skilled in the art should appreciate that an appropriate frame size can
30 be established from an image having more or fewer pixels, by redundantly allocating data into adjacent pixels or by eliminating intermediate pixels, or similar technique.

In another aspect, a camera is provided which optically "zooms" to provide optimal imaging for a desired image zone. By way of example, the invention of one aspect takes an image of the user's head, determines the location of the user's eyes
5 (such as by matched filtering), and optically zooms the image through movement of optics to provide an image of the eyes in the desired processing size format.

Many aspects of the invention are preferably enhanced by autofocus. Specifically, it is often desirable to have a crisp image of the user (or a part of the
10 user, e.g., the user's eyes) in order to accurately process desired cursor or scene view movement. Thus, autofocus capability preferably operates in most of the aspects of the invention where imaging is a feature of the processing.

In one aspect, the camera utilizes a very small aperture which results in a very large depth of field. In such a situation, autofocus is not required or desired. The
15 optical requirements for the lenses are also reduced.

The invention thus provides several advantages over the art. For example, game controllers can now include feedback corresponding to the user's actual movement. By way of another example, if the user moves left or right (or head or
20 hand or eyes move left or right, depending on the image zone), then the cursor (or scene view) can also be set to move left or right. When the user twists her head, for example, the scene view can also be made to rotate, reflecting that movement.

Those skilled in the art should appreciate that the direction in which the scene
25 moves, left or right, is a matter of design choice. That is, certain games might find it desirable to move the opposite direction from what the user moves, to add certain challenges to the game. Further, in other aspects, this direction can change during the game to further complicate game control.

30 In accord with one aspect of the invention, a processing subsystem (connected with the camera) is used to make cursor movement (or scene view movement)

correspond to user's motion. This processing subsystem of another aspect further detects when the user twists his head, to add an additional dimension to the movement.

5 In one aspect, a system of the invention includes an IR detector which is used to determine when a person sweats or heats up (by imaging, for example, part of the user's head onto the IR detector); and then the system adjusts game speed in a way corresponding to this movement. Alternatively, a heartbeat sensor is tied to the person to sense increased excitement during a game and the system speeds or slows
10 the game in a similar manner. Note that a heartbeat sensor can be constructed, in one aspect of the invention, by thermal imaging of the user's face, detecting blood flow oscillations indicative of heartbeat. In other aspects, the heartbeat sensor is physically tied to the user, such as within the computer mouse or joystick.

15 In one aspect, a computer of the invention adapts to user control as selected by a particular user. For example, in the case of a handicapped person, a particular user might select certain hand-movements, e.g., a single finger up, to move the cursor up; and another finger down to move the cursor left. An infinite combination of controls can be established; however this is one advantage of the invention in that
20 users with many different disabilities can program cursor or scene view movement. In one aspect, a neural network is used to assist the processing system in establishing proper cursor movement. In another aspect, the computer for example learns to print something by movement of the user's finger (or other body part).

25 In one aspect, tipping of the user's head (or other body part, or object) is used to provide another degree of freedom in moving the cursor or adjusting the scene view. By way of example, a tilt of the head, as imaged by the camera, can be set to command a rotation of the scene view.

30 In still another aspect, a camera of the invention uses autozoom to move in and out of a given scene view. By way of example, the camera is first focussed on the

user's face in one frame; but in a subsequent frame the camera must focus to closer to compensate for the fact that the user moved closer to the camera (typically, the camera is on the monitor, so this also means that the user moved closer to the scene view). This autofzoom is used, in one aspect, to make the scene view appear as if the user is "creeping" into the scene. By moving the scene in and out, the user will perceive that he is moving in or out of the scene view.

In another aspect, a camera images an object held by the user. Preferably, the object has a well-defined shape. The system images the object and determines difference information corresponding to movement of the object. By way of example, rotating the object upside down results in difference information that is upside down; and then the scene view inverts by operation of the system. In another example, twisting of the object rotates the scene view left or right, or rotates the scene in the direction of the twisting.

In another aspect, two cameras image the user: one camera pointed at the front of the users face or hand and the other down at the top of the users head or hand. The front facing camera is used to detect rotational and linear translation in up-down and left-right directions. The top viewing camera determined front-back, left right translation. The front-back translation observed by the top camera is used to control forward and back motion in the users 3-D view. The top sensed left-right translation controls the users left right slide or strafe. The top sensed left-right motion is removed from the front view left-right translation with the remaining front view measure representative of left-right twist. All of the front view up-down translation can be interpreted as up-down twist.

Brief Description of the Drawings

Figure 1 illustrates one human computer interface system constructed according to the invention;

Figure 1A shows an exemplary computer display illustrating cursor movement made through the system of Figure 1;

5 Figure 1B illustrates overlaid scene views, displayed in two moments of time on the display in Figure 1, of a shifting scene made in response to user movement captured by the camera of Figure 1;

Figure 1C shows an illustrative frame of data taken by the system of Figure 1;

10 Figure 2 illustrates selected functions for a printed circuit card used in the system of Figure 1;

Figure 3 illustrates an algorithm block diagram that preferably operates with the system of Figure 1;

15

Figure 4 illustrates one preferred algorithm process used in accord with the invention to determine and quantify body motion;

Figure 5 shows one process of the invention for communicating body motion data to a host processor, in accord with the invention, for augmented control of cursor position or scene view;

20

Figure 5A shows representative frame of data of a user taken by a camera of the invention, and further illustrates adding symbols to key body parts to facilitate processing;

25

Figure 6 illustrates a two camera imaging system for implementing the teachings of the invention;

Figure 7 illustrates two positions of a user as captured by a camera of the invention; and Figure 7A illustrates two positions of a scene view on a display as repositioned in response to movement of the user illustrated in Figure 7;

5 Figure 8 illustrates motion of a user - and specifically twisting of the user's head - as captured by a system of the invention; Figure 8A illustrates a first scene view corresponding to a representative computer display before the twisting; Figure 8B illustrates a second scene view corresponding to a rotation of the first scene view in response to the twisting by the user; Figure 8C shows processing features of the
10 processing section of Figure 8; and Figure 8D illustrates multiple image frames stored in memory for matched filtering with raw images acquired by the system of Figure 8;

Figure 9 illustrates a two camera system of the invention for collecting N
15 zones of user movement and for repositioning the cursor or scene view as a function of the N movements; Figure 9A illustrates a representative thermal image captured by the system of Figure 9; and Figure 9B illustrates process methodology for processing thermal images as a real time input to game processing speed, in accord with the invention;

20

Figure 10 illustrates another two camera system of the invention for targeting multiple image movement zones on a user, and further illustrating optional DSP processing at the camera section;

25 Figure 11 illustrates framing multiple movement zones with a single imaging array, in accord with the invention;

Figure 12 illustrates framing a user's eyes in accord with the invention; and Figure 12A shows a representative image frame of a user's eyes;

30

Figure 13 illustrates one system of the invention, including zoom, neural nets, and autofocus to facilitate image capture;

Figures 14, 14A and 14B illustrate autofocus motion control in accord with the invention;

Figures 15 and 15A illustrate one other motion detect system algorithm utilizing edge detection, in accord with the invention;

Figure 16 illustrates one other motion detect system algorithm utilizing well-characterized object manipulations, in accord with the invention;

Figure 17 illustrates one other motion detect system algorithm utilizing varied body motions, in accord with the invention;

Figure 18 illustrates a two camera system of the invention with a camera observing the user's face while the other observes the top of the user's head;

Figure 19 shows a blink detect system of the invention; and

Figure 20 shows a re-calibration system constructed according to the invention.

Detailed Description of the Drawings

Figure 1 illustrates, in a top view, certain major components of a human computer interface system 10 of the invention. A user 12 of the system 10 sits facing a computer monitor 14 with display 14a. A camera 16 is mounted on the computer monitor 14 facing the user 12. In the illustrated embodiment, the camera 16 is mounted in such a way that the user's face 12a is imaged within the camera's field of view 16a. However, as discussed herein, the camera 16 can alternatively image other

locations, such as the user's hand, eyes, or on other objects; so imaging of the user's face, in Figure 1, must be considered illustrative, rather than limiting. Further, the camera location can also reside at places other than on top of the monitor 14.

5 With further regard to Figure 1, the camera 16 interfaces with a printed circuit card 18 mounted within the user's computer chassis 20 (which connects with the monitor 14 by common cabling 20a). The camera 16 interfaces to the printed circuit card 18 via a camera interface cable 22. The circuit card 18 also has processing section 18a, such as a digital signal processing ("DSP") chip and software, to process
10 images from the camera 16.

 In operation, the camera 16 and card 18 capture frames of image data corresponding to user movement 25. The processing section 18a algorithmically processes the image data to quantify that movement 25; and then communicates this
15 information to the host processor 30 within the computer 20. The host processor 30 then commands movement of the computer cursor in a corresponding movement 25a, Figure 1A (Figure 1A illustrates a representative front view of the display 14a, and also illustrates movement 25a of the cursor 26 moving within the display 14a in response to user movement 25).

20 Figure 1B illustrates an alternative (or supplemental) process whereby the scene view shifts in response to user movement 25. Specifically, Figure 1B illustrates a first scene view 35a, which generally corresponds to a forest prior to the user's movement 25; and an overlaid scene view 35b (shown in dotted line, for purposes
25 of illustration) that is shifted by an amount 37 in response to the user's movement 25. The shift 37 in the scene view 35 is accomplished by combined operation and processing of the processing section 18a and host CPU 30.

 Figure 1C shows a representative frame 41 of data 43 as taken by the camera
30 16. As illustrated, data 43 represents the user's face 12a taken at a given moment of time. Subsequent frames (not shown) are used to determine user motion 25 relative

to the frame 41, as discussed herein. The frame 41 is made up of the plurality of pixel data 45, as known in the art.

Figure 2 illustrates certain functions processed within the printed circuit board 18 of Figure 1. A camera interface circuit 50 receives video data from the camera 16 through interface cable 22. This video data can be RS170 format or digital, for example. For analog RS170 format, circuit 50 decodes the analog video data to determine video timing signals embedded in the analog data. These timing signals are used for control of the analog-to-digital (A/D) converter included in circuit 50 that converts analog pixel data into digital images. In the preferred embodiment, the analog data is digitized into 6-bits, though any number of bits greater may be acceptable and/or required for features as discussed herein. For digital data format, camera interface 50 accepts the digital data without additional quantization, although interface 50 can digitally pre-process the digital images if desired to acquire desired image features.

The frame difference electronics 52 receives digital data from the camera interface circuit 50. The frame difference electronics 52 include a multiple frame memory, a subtraction circuit and a state machine controller/memory addresser to control data flow. The frame memory holds previously digitized frame data. As each digitized pixel is received by the frame difference electronics 52, the corresponding pixel from a previous frame is read from the frame memory and subtracted from the current frame. The preferred implementation uses the frame just previous to the current frame, though an older frame which resides in the frame memory may be used. The resulting difference is output to an N-frame video memory 54. The new frame pixel data is then stored into the frame memory of the frame difference electronics 52.

The N frame video memory electronics 54 either receives differenced frames output by the frame difference electronics 52 (discussed above) or raw digitized frames from the camera interface 50. The choice of where the data derives from is

made by software resident on the DSP 56. The frame video memory 54 is sized to hold more than one full frame of video and up to N number of frames. The number of frames N is to be driven by hardware and software design.

5 In the preferred embodiment, the DSP 56 implements an algorithm discussed below. This algorithm determines the rate of head motion of the user in two dimensions. The digital signal processor 56 also detects the eye blink of the user in order to emulate the click and double click action of a standard mouse button. In support of these functions, the DSP 56 commands the N frame video memory 54 to
10 supply either the differenced frames or the raw digitized frames. The digital signal processor thus preferably utilizes a supporting program memory 58 made up of electrically reprogrammable memory (EPROM) and data memory 59 including standard volatile random access memory (RAM). The DSP 56 also interfaces to the PCI bus interface electronics 60 through which cursor and button emulation is
15 passed to the user's main processor (e.g., the CPU 30, Figure 1). The PCI interface 60 also passes raw digitized video to the main processor as an optional feature. Interface 60 also permits reprogramming of program memory 58, to allow for future software upgrades permitting additional features and performance.

20 The PCI interface electronics 60 thus provides an industry standard bus interface supporting the aforementioned communication path between the printed circuit card 18 and the user's main processor 30.

With optional MPEG compression electronics 62, the printed circuit card 18
25 and camera 16 can provide compressed video to the user's main processor 30. This compressed video supports using the system 10 in teleconferencing applications, providing dual use as either human computer interface system 10 and/or as a teleconferencing system in an economical solution to two distinct applications.

30 Figure 3 describes one preferred head motion block diagram algorithm 70 used in accord with the invention. Not all of the functions shown in Figure 3 are

implemented in software in the DSP 56. Rather, this algorithm relies on the correlation of images from one frame to the next, and particularly relies on the use of frame differenced images in the correlation process. The frame differencing operation removes parts of the camera images that are unchanged from the previous frame. For example, room background (such as object 13, Figure 1) behind the user 12 is removed from the image. This greatly simplifies detection of feature motion. Even the image of the user's face image consists of regions of uniform illumination such that even with the user's facial motion, these uniform regions (i.e. cheeks, forehead, chin) may also be removed. The user's face 12a also consists of typically dynamic features such as the nose, eyes, eyebrows and mouth, each of which typically has enough spatial detail that will be evident in the differenced image. As the user moves his face with respect to room lighting, the shape and distribution of these features will change; but the frame rate of the camera 16 ensures that these features look similar from one frame to the next. The correlation process therefore operates to determine how these differenced features are moving from one frame to the next in order to determine user head motion 25.

The algorithm of block diagram 70, Figure 3, receives video images 72 of the user as imaged by camera 16 over time. Each received image is passed to both a frame memory 74 and a differencer 76. Though the preferred embodiment is to buffer a single frame in memory 74, the memory 74 may optionally store many frames, buffered such that the first frame input is the first frame output. The delayed frame is read from the frame memory 74 and subtracted from the current frame using the differencer 76. Frame output from the differencer 76 is provided to both a correlation process 78 and a difference frame memory 80.

Like the frame memory 74, the preferred embodiment of frame memory 80 utilizes a single difference frame; however the difference frame memory 80 can hold many difference frames in sequence for a finite time period. The delayed difference frame is read from the difference frame memory and provided to the correlation function 78. Difference frames are preferably selectable by system algorithms. The

correlation process 78 determines the best combination of row and column shifts in order to minimize the difference between the current difference frame and the delayed difference frame. The number of rows and columns required to align these difference images provides information as to the user's motion.

5

The best-fit function algorithm 82 determines the row and column shift to provide optimal alignment. In the case of a classical correlation process, the best-fit function can consist of a peak detect algorithm. This algorithm may either be implemented in hardware or in software.

10

The best-fit function algorithm determines relative motion in rows and columns of the observed user's features. The cursor update compute function algorithm 84 translates this measured motion into the position change required of the cursor (e.g., the cursor 26, Figure 1A). Typically, this is a non-linear process that, with greater head motion, the cursor moves a non-proportionally greater distance. For example a 1-pixel user motion can cause the cursor to move one screen pixel while a 10-pixel user motion may cause a 100-pixel screen cursor motion. However, these magnifications can be adjusted for desired result. This algorithm may either be implemented in hardware or in software such as through an ASIC or FPGA.

20

Video cursor control 86 provides a user interface to enable and disable the operation of cursor control described above. This control is implemented, for example, through a combination of keystrokes on the user's keyboard (for example as connected to the host computer 20, Figure 1). Alternatively, cursor control is activated or deactivated by sensing the eye-blink of the user (or some other predetermined movement). In this alternative embodiment, an output signal 85 from the correlation section 78 is sent to the video enable section 86; and the output signal 85 corresponds to blink data from the user's face 12a (Figure 1A). In another embodiment, the video cursor control section 86 activates or deactivates cursor control by recognizing voice commands. A microphone 87 detects the user's voice and a voice recognition section 89 converts the voice to certain activate or deactivate

30

signals. For example, the section 89 can be set to respond to "activate" as a voice command that will enable cursor control; and "deactivate" as a command that disables cursor control.

5 The functionality of the video cursor control 86 provides the user with the equivalent of a mouse pick-up, put-down action. As the user moves the cursor from left to right across the screen, the user would de-activate motion-based cursor control in order to allow the user to move his head back to the left. Once the user has recentered his head, the user would once again activate the cursor control and
10 continue to move the cursor about the screen. The activation/deactivation of the mouse input is represented by the switch 90, such that the open position of the switch disables human motion control of the cursor and supplies a zero change input to the summation operation 92 in such conditions.

15 Those skilled in the art should appreciate that control of scene view may also be implemented by an algorithm such as shown in Figure 3. Specifically, a similar algorithm can provide movement of the current scene view, in accord with the invention.

20 With video cursor control enabled, the result of the cursor update compute function 84 is added to the known current cursor position at the summing operation 92. This summation has a x component and a y component. The result of the summation 92 is used to update the cursor position (or scene view) on the user's screen via the user's operating system. Cursor position may thus be controlled by
25 both user motion as well as the motion imparted by another input device such as a standard computer mouse.

Figure 4 provides a detailed description of the preferred implementation of the algorithm described in functions 73, 76, 78, 80 and 82 of Figure 3. Video data is
30 received by the processing electronics in both a single frame memory 100 and a differencer 102. The output of the frame memory 100 is also provided to the

differencer 102 such that the previous frame is subtracted from the current frame. This differenced frame is then processed by a two dimensional FFT 104. The complex result of the FFT 104 is provided to a complex multiplier 106 and a complex memory 108. The complex memory 108 is the size of the processed image, each location
5 containing both a real and imaginary component of a complex number. With each new FFT operation 104, the previous FFT result, contained in the complex memory 108, is provided to the conjugate operation 110. The complex conjugate of each element is computed and provided to the complex multiplier 106. In this manner, the FFT of the previous frame difference is conjugated and multiplied against the
10 FFT of the current difference image.

By way of comparison between Figures 3 and 4, item 76 has similar functionality to item 102; item 78 has similar functionality to items 104, 106, 108, 110, 112; item 80 has similar functionality to item 108; and item 82 has similar
15 functionality to item 114.

The two dimensional array of complex products output by the complex multiplier 106 is provided to a two dimensional inverse FFT operation 112. This operation creates an image of the correlation function 114 between the latest pair of
20 difference images. The correlation image is processed by the peak detection function 114 in order to determine the shift required in aligning the two difference images. The x-y magnitude of this shift is representative of the user's motion. This x-y magnitude is provided to the software used to update the cursor position as described in Figure 3.

25

Figure 5 shows an algorithm process 130 of the invention and which applies motion correlation operations over sub-frames of the video image. This allows motions of various body parts to convey input with specialized meaning to applications operating on the host computer. In addition to head motion, motion of
30 the hands, arms and legs provide for greater degrees of freedom for the user to interact with the host application (e.g., a game). Commands of this type are useful in

combative games where computer animated opponents fight under control of the user. In that instance, the hand, arm and leg motions of the user become punch, chop and kick commands to the computer after process 130. This command mode can also be used in situations where the user does not have ready access to a keyboard,
5 to augment cursor control of the previously described head position correlator.

Process 130 identifies the functions required to derive commands from general motions of the user's body. The scene analyzer function 132 receives digitized video frames from the camera (e.g., the camera 16 of Figure 1) and
10 identifies sub-frames within the video for tracking various parts of the user's body. The frame difference function 134 and correlator function 136 provide similar functions as processes 74, 76 and 78 of Figure 3. The correlation analyzer 138 receives correlated difference frames from the correlator function 136 and sub-frame definitions from the scene analyzer 132. The correlation analyzer 138 applies a peak
15 detection function to each sub-frame to identify the shift required to achieve best alignment of the two images. Correlation peaks occurring in the center of the sub-frame indicate no motion, while peaks occurring elsewhere indicate the direction and magnitude of the user's motion. The motion interpreter 140 receives motion vectors for each sub-frame from the correlation analyzer 138. The motion interpreter
20 140 links the motion vector from each sub-frame with a particular body segment and passes this information onto the host interface 142. The host interface 142 provides for communication with the host processor (e.g., CPU 30, Figure 1). It sends data packets to the host to identify detected body motions, their directions and their amplitudes. The host interface 142 also receives instruction from the host as to which
25 body segments to track which it then passed along to the motion interpreter 140 and the scene analyzer 132.

The scene analyzer 132 first identifies the location of the user's body in the image and locates the position of various parts of the user's body such as hands,
30 forearms, head, and legs. The techniques and methods used to identify the user's body location and body part positions can be accomplished using techniques well

known to those skilled in the art (by way of example, via matched filtering). Body identification can also be augmented by marking different locations on the user's body with unique visual symbols. Unique symbols are assigned to key body joints such as elbows, shoulders, hands, neck, knees, and waist and are mounted on the body. See for example Figure 5A.

Figure 5A illustrates one frame 149 of data of an image of the user 150 as taken by a camera of the invention. The image corresponds to a full body image of the user 150, including arms 151, legs 152, elbows 151a, hands 153, head 154, neck 155, ears 156, and forehead 157. These parts 151-157 are identified by processes of the invention (e.g., spatial location in the image, by matched filtering or other image recognition technique), and the image is preferably marked with unique symbols (e.g., "X" for center of the face, "Y" for center of the hand 153, "T" for center of the user's foot, "Z" for body center, and "F" for forehead 157).

With further reference to Figure 5, process 130 locates various body parts and preferably marks them with symbols to fill in connecting logic (e.g., the left wrist and left elbow symbol identify the location of the left forearm). Once the user's body parts are located, sub-frames surrounding each of the body segments identified by the host processor are generated. A sub-frame is a generally regularly shaped region within the image that surrounds a particular body part. The sub-frames are sized to center the subject body part in the sub-frame and to provide enough room around the body part to accommodate typical body motions. One sub-frame 160 is shown in frame 149, Figure 5A, surrounding the user's foot "T". The scene analyzer 132 will generally not operate on each frame of video since continuously changing the sub-frames adds unnecessary complication to the correlation analyzer 138. Instead, the scene analyzer 132 runs as a background process updating the sub-frame locations periodically.

Figure 4 provides a detailed description of one algorithm which can be used to implement processes 134-138 of Figure 5.

The invention of one embodiment can thus track the motion of the user's body using symbols attached to key joints. As an example, the position of the user's left lower arm can be determined by locating the unique symbol for the left hand
5 "Y1" and left elbow "●". Unique symbols thus allow the processor to rapidly locate each portion of the user's body in a video frame. To determine the motion of a particular part of the user's body, the algorithm (e.g., Figure 4) compares the position of the relevant body parts in consecutive frames and determines how they moved (for example, using geometry). Once motion is determined, it is then passed
10 to the host CPU where the motion is acted on as appropriate for the particular application.

Figure 6 illustrates a two camera system 200, constructed according to the invention. The cameras 202a, 202b are arranged to view separate parts of the user:
15 camera 202a images the user's face 204; and camera 202b images the user's hand 206. The cameras 202 conveniently rest on top of the computer display 208 coupled to the host computer 210 by cabling 216. The cameras 202 couple to the signal processing card 212 residing within the computer 210 by cabling 213. As discussed herein, motion of the user's head 204 and/or hand 206 are detected by the signal processing
20 card 212, and difference information is communicated to the computer's CPU 210a via the computer bus 214. This difference information corresponds to composite movement of the head 204 and hand 206; and is used by the CPU 210a to command movement of display items on the display 208 (for example, the display items can include the cursor or scene view as shown on the display 208 to the user).
25 Information shown on the display 208 is communicated from the computer 210 to the display 208 along standard cable 216.

Figures 7 and 7A illustrate how motion of a user's head is for example translated to motion of the cursor and/or scene view, in accord with the invention.
30 Figure 7 shows a representative image 220 of a user captured within a frame of data

by a camera of the invention. Figure 7 also shows a representative image 222 (in dotted outline, for clarity of illustration) of the user in a subsequent frame of data, indicating that the user moved "M" inches. Figure 7A illustrates corresponding scene views on a computer display 224 that is coupled to processing algorithms of the invention (i.e., within a system that includes a camera that captures the images 220, 222 of Figure 7). The display 224 illustratively shows a scene view that includes a road 224a that extends off into the distance, and a house 224b adjacent to the road 224a. A computer cursor 224c is also illustratively shown on the display 224 as such a cursor is common even within computer games, providing a place for the user to select items (such as the road or house 224a, 224b) within the display 224. The display 224 also shows, with dotted outlines 226, the scene view of road and house which are shown on the display 224 after motion by the user from 220 to 222, Figure 7 (the cursor 224c is for example repositioned to position 224c'). The repositioning of the scene view from 224a, 224b to 226 occurs immediately (typically less than 1/30 second, depending upon the camera) after the movement of the user of Figure 7 from 220 to 222. The scene view is repositioned by x-pixels on the display 224, so that M/x corresponds to the magnification between user movement and scene view repositioning. This magnification can be set by parameters within the system; and can also be set by the user, if desired, at the computer keyboard. Furthermore, the rate at which the scene view moves the distance of x-pixels preferably occurs at the same rate as the rate of travel along distance M. Alternatively, the magnification can be dependent on the rate of motion such that a larger displacement of x-pixels will occur for a given motion M if the rate of change of M is larger.

Figure 8 illustrates a further motion that can be captured by a camera of the invention and processed to reposition a scene view, as shown in Figures 8A and 8B. More particularly, Figure 8 illustrates a camera 250 connected to a processing section 252 which converts user motion 254 to corresponding repositioning of the computer scene view. As above, the user 256 is captured by the camera's field of view 258 and frames of data are captured by the processing section 256. In Figure 8, motion 254 corresponds to a twisting of the user's head 256; and processing section 252 detects

this twisting and provides repositioning information to the host computer (not shown). Processing section 252 can also incorporate head-translation motion (e.g., illustrated in Figure 15A) into the scene view movement above; and can similarly reject translational movement too, if desired, so that no scene motion is observed for translation of the user 256.

Figure 8A shows a representative scene view 260 on a display 262 coupled to the host computer. Figure 8B illustrates repositioning of the scene view 260' after the processing section 252 detects motion 254 and updates the host computer with difference information (e.g., that information which the host computer uses to rotate or translate the scene view).

Figure 8A also illustrates the intent of the rotating scene view feature. In Figure 8A, a person 260a is shown in the scene view 260, except that the person 260a is almost completely obscured by the edge 262a of the display 262. By twisting the head 256 in motion 254, the scene view 262 is rotated in the corresponding direction - as shown by scene view 260' in Figure 8B - so that the user 260a' is completely visible within the scene view 260'.

Figure 8C illustrates further detail of the processing section 252. Camera data such as frames of images of a user are input to the section 252 at data port 266. The data are conditioned in the image conditioning section 268 (for example, to reduce correlated noise or other image artifacts). Thereafter, the camera data is compared and correlated in the image correlation section 270, which compares the present frame image with a series of stored images from the image memory 272. In the preferred embodiment, the present data image frame 249 is cross-correlated with each of the images within the image memory 272 to find a match. These images correspond to a series of images of the user in known positions, as illustrated in Figure 8D.

In Figure 8D, various images are stored representing various known positions of relevant part, here the user's head 256. In the position of Figure 8, for example a straight on face shot, the 0° stored memory image would provide the greatest cross-correlation value indicating a matched image position. Accordingly, the scene view
5 would adjust to a zero position. If, however, the image correlated to a -90° position, the scene would rotate to such a position. Other movements cause additional scene view motions, including tilt and tip of the head, as shown in the two images "0°, Down 45°" image and the "0°, Up 45°". These images cause the scene view to move upwards or to tilt up or down, when the process section 252 correlates the current
10 frame to these images. As indicated, these images have no left or right component, though other images (not shown) can certainly include left, right and tip motion simultaneously.

Figure 9 shows a system 300 constructed according to the invention and
15 including a camera section 302 including an IR imager 304 and a camera 306, both of which view and capture frames of data from a user 308. The IR imager 304 can include, for example, a microbolometer array (i.e., "uncooled" detectors known in the art) which produces a frame of data corresponding to the infrared energy emitted from the user, such as illustrated in Figure 9A. Figure 9A shows a
20 representative frame of IR image data 310, with zones 312 of relatively hot image data emitting from regions of forehead, nose and mouth of the user 308.

The cameras 304, 306 send image data back to the signal processing section 314. Data from the camera 306 is processed, if desired, as above, to determine
25 difference information signal 322 used by a connected computer to reposition the cursor and/or scene view. Data from camera 304, on the other hand, is used to evaluate how much (or how hot) zones 312 appear on the user during play of the computer. The signal processing section 314 assesses the zones 312 for temperature and/or size over the course of a computer game and generate a "game speed
30 control" signal 320 which is communicated to the user's computer (i.e., that computer used in conjunction with the system 300 of Figure 9). The user's computer

processes the signal 320 to increase or decrease the speed of a computer game in process on the computer.

Those skilled in the art should appreciate that the IR camera 304 can be used
5 without the features of the invention which assess user movement. Rather, this aspect should be considered stand-alone, if desired, to provide active feedback into gaming speed based upon user temperature and/or stress. Note that the camera 304 can also be used to detect heartbeat since the zones 312 generally pulse at the user's heartbeat, so that heartbeat rate can also be considered as a parameter used in the
10 generation of the game speed control signal 320. Alternatively, a pulse rate can be determined by known pulse rate systems that are physically connected to the user 308.

An IR lamp 324 can be used in system 300 to illuminate the user 308, with IR
15 radiation 324a, such that sufficient IR illumination reflects off of the user 308 whereby motion control of the cursor and/or scene view can be made without the additional camera 306. The lamp 324 can be, and preferably is, made integrally with the section 302 to facilitate production packaging.

20 An IR lamp 324 operating in the near-IR can also be used with visible cameras of the invention which typically respond to near-IR wavelengths. By way of example, certain camera systems now available incorporate six IR emitters around the lens to illuminate the object without distraction to the user who cannot see the near-IR emitted light. Such a camera is suitable for use with the invention.

25

Figure 9B shows process methodology of the invention to process thermal user images in accord with the preferred embodiment of the invention. Specifically, a system such as system 300 first acquires a thermal image map in process block 326. This image is compared to a reference image ("REF") in process block 327. REF can
30 either be a temperature of the user (i.e., a temperature of one hot spot of a non-stressed user, or the temperature of one hot spot of the user at an initial, pre-game

condition) or an amount of the area 312, Figure 9A, of the user in a non-stressed condition or initial pre-game condition). By way of example, REF can be an image such as the frame 310 of Figure 9A. When the temperature or area of the region 312 increases, the system 300 detects this change and determines that the image map
5 exceeded the REF condition, as illustrated in process block 328. Should the map exceed the REF condition, the system 300 communicates this to the host processor which in turn adjusts the gaming speed, as desired. If the map does not exceed the REF condition, then the next IR image frame is acquired at block 326.

10 System 300 and the process steps of Figure 9B are thus suitable to adjust gaming speed in real time, depending upon user stress level. In the preferred embodiment, the gaming speed is increased automatically such that the image map exceeds the REF signal for greater than about 50% of the time, so that all users, regardless of their ability, are pushed in the particular game.

15 Those skilled in the art should appreciate that multi-camera embodiments of the invention can and preferably are incorporated into a common housing 338, such as shown in Figure 10. Further, as illustrated in Figure 10, cameras can also be made from detector arrays 340, processing electronics 342, and optics 344. Each camera
20 340, 342, 344 is constructed to process the correct electromagnetic spectrum, e.g., IR (using, for example, germanium lenses 344 and microbolometer detectors 340). Each camera has its own field of view 350a, 350b and focal distance 352a, 352b to image at least a part of the user. These field of views 350 can overlap, to view the same area such as the user's face, or they can view separate locations, such as the user's head
25 and hand.

Cameras of the invention can also include a DSP section 356 such as described above to process user motion data. The DSP section 356 processes user motion data and sends difference information to the user's host computer. The host computer
30 thereafter repositions the cursor and/or scene view based upon the difference information so that the user observes corresponding motion on the computer

display, as described above. Accordingly, the DSP section need not reside within the computer so long as difference information is isolated and communicated to the host computer CPU.

- 5 Those skilled in the art should appreciate that algorithms of the invention can also be processed directly by the computer's CPU, provided it has sufficient power and processing speed, to eliminate a separate DSP chip or section. DSP or equivalent processing capability can also be provided within the computer's chassis by way of a computer printed circuit card installed in the chassis and connected with the camera.
- 10 The location and amount of processing power, therefore, should be considered a matter of design choice and current state of the art, each technique being within the scope of the invention.

- Figure 11 illustrates frame capture by one camera of the invention to isolate
- 15 zones of imaging according to expected motion patterns. In Figure 11, one frame 370 of data for example covers the user's eyes 371, corresponding to one image zone; and another frame 372 of data can cover the user's head 373, corresponding to another image zone. As mentioned previously, preferably the frames 370, 372 are 64x64 pixels each, or 256x256 (or higher powers of two) to provide FFT capability on the image within the frame. A single camera can however provide both frames 370 and
- 20 372, in accord with the invention. Specifically, a dense CCD detector array (e.g., 480x740 pixels, 1000x1000 pixels, or higher) is used within the camera such that the whole array captures an image frame 376 of data, at least covering the available image format of the computer display 378. A matched filtering (or other image
- 25 locate process) is processed on the frame 376 to locate the center 371a of the user's eyes (in the matched filtering process, an image data set of the user's eyes is stored in memory and correlated to the frame 376 such that a peak correlation is found at position 371a). Thereafter, a 64x64 array of data is centered about the eyes 371 to set the frame 370. To acquire the frame 372, every other pixel is discarded so that, again,
- 30 a 64x64 array is set for the frame 372 (alternatively, each adjacent pair of pixels is added and averaged to provide a single number, again reducing the total number of

pixels to 64x64). Note that this process is reasonable since the width of the eyes is at least $\frac{1}{2}$ the width of the user's face. Nevertheless, further compression can be obtained by utilizing every third pixel (or averaging three adjacent pixels) to obtain a larger image area in the frame 372. Note that the compression in the width and
5 length dimensions need not be the same.

Framing of the information in Figure 11 can occur in several ways. Most cameras image at 30Hz so that image motion is smooth to the human eye. In one embodiment, one frame 370 is taken in between each frame 372, to minimize data
10 throughput and processing; and yet to maintain dual processing of the two zones imaged in Figure 11. Alternatively, both frames 370, 372 are processed concurrently since frame 376 is typically the 30Hz frame.

Figure 11 also illustrates how framing can occur around the user's eyes 371 to
15 acquire "blink" information to reset cursor control. A blink detected by the user's eyes in frame 370 (or other frame) can be used to (a) disable or enable control of cursor or scene movement based upon user control, or (b) simulate pick-up and replacement of the computer mouse (i.e., reinitializing movement in a particular direction). For example, by detecting a blink of the eyes 371, a system of the
20 invention can disable human motion following control such as described herein. Another blink can be used to enable human motion following control. Blinking can also be used to continue motion in a particular direction. For example, movement of the cursor can be made to follow movement of the user's head, as described above. However, after a while, the person has to move to an uncomfortable position to keep
25 moving the cursor (or scene). A blink can thus also serve to reposition the head back to a normal starting position so that further movement in the desired direction can be made.

Figure 12 illustrates a similar capture of a user's eyes 400, in accord with the
30 invention. A frame 402 can thus be acquired by a camera of the invention. Figure 12A illustrates further detail of one representative frame 402, illustrating that the

user's pupils 404 are also captured. Figures 3 and 4 describe certain algorithms of the invention that are also applicable to motion of the user's pupils 404, as illustrated by left and right motion 406 and up and down motion 408. Accordingly, by zooming in on the user's eyes, another movement zone is created that causes repositioning of the cursor or the scene view based upon the movements 406, 408, much like the head movement described and illustrated in Figures 1-4.

Note that the teachings of Figures 1-4 and 12-12A can be combined within a two zone movement system so that, for example, both head motion and pupil motion can be evaluated for image motion. The cursor and/or scene view can be repositioned, therefore, based upon movements from both zones. By way of example, repositioning of items within the display (e.g., the cursor and/or scene view) can be made when the head moves but not if the head and eyes move, which might indicate that the user is simply looking elsewhere in the room due to a distraction. However, if the user moves his head, but not his eyes, he is focussed on the game and intends rotation of the scene view, in another example. Other combinations are also possible.

Cameras of the invention can also include zoom optics which (a) reduce or enlarge the image frame captured by a particular camera, or which (b) provide autofocus capability. Figure 13 shows one system 430 constructed according to the invention. A camera 432 includes camera electronics 432a and a zoom attachment 432b. Data from the camera 432 is relayed to image and interpretation feedback electronics 434 for evaluation. For purposes of image magnification, the feedback electronics serve to evaluate a given image size relative to desired image goals. For example, to image the user's eyes with high fidelity might require high density of pixels at the user's eyes (e.g., at the zone 370, Figure 11). Accordingly, the system 430 can isolate the user's eyes, such as described herein, and command the camera (through command lines 436) to increase or decrease magnification on the user's eyes so as to achieve desired resolution. The feedback electronics can also command motion of the camera to change its boresight alignment (i.e., to change where the

camera image is centered) by commanding movement of the camera when resting on one or more linear drives 438, as known in the art.

Once aligned on the desired user location, e.g., on the eyes with desired accuracy, the system 430 continues processing data such as described herein to create human interface control of items displayed on the user's computer, e.g., cursor and/or scene view. Accordingly, processing section 440 operates to detect user motion and to communicate difference information to the user's computer, as described above.

With regard to autofocus, the system 430 of Figure 13 can also be used to process user motion based upon motion towards and away from the camera. Figure 14 illustrates such a system, including a camera 450 with autofocus capability to find the best focus 452 relative to a user 454 within the field of view 456. For example, when the user 454 moves to position 460 (the user being shown in outline form 454a), the new best focus has changed to 452a. The camera 450 provides a signal 450a to the image interpretation and feedback electronics 434, Figure 13, which indicates where the user is along the "z" axis from the camera 450 to the user 454. This signal 450a is thus used much like the other motion signals described herein, to move the cursor and/or scene view in response to such movements. Figure 14A illustrates a representative scene view 462 when, for example, the user is at best focus 452. The scene view 462 includes a house image 464 with a door 465. When the user moves to position 460, the house and door 464', 465' of the scene view 462' enlarge, since the user moved closer to the camera 450. Such a motion might reveal, for example, additional objects within the house, such as illustrated by object 466, Figure 14B. Accordingly, the autofocus feature of the invention provides yet another degree of freedom in motion control, in accord with the invention.

Image data, manipulation, and human interface control can be improved, over time, by using neural net algorithms. As shown in Figure 13, a neural net update section 435 can for example couple to the feedback electronics 434 so as to

assimilate movement information and to improve data transmitted to the host computer, over time. Use of neural nets are known in the art.

Figure 15 illustrates a frame of data 490 used in accord with the invention to
5 implement a simplified left, right, up, down movement algorithm to control cursor
movement and/or scene view movement. Frame 490 is captured by a camera of the
invention; and preferably the camera incorporates autofocus, as described above, to
provide a crisp image of the user 492 regardless of her position within the camera's
field of view. As shown, image frame 490 provides very sharp edges to the user's
10 face, including a left edge 494a, right edge 494b, and chin 494c. These edges need
only approximate vertical or horizontal position. Movement of the user results in
movement of the edges 494, such as shown in Figure 15A. Figure 15A shows that
once such edges are acquired, they conveniently permit subsequent movement
analysis and control of scene view and/or cursor position. Specifically, Figure 15A
15 shows movement of the user's "edges" from 494a-c to 494a-c', indicating that the
user moved left (as viewed from the camera's position) and that her chin raised
slightly, indicating that an upward tilt of the head. This information is assessed by
the process sections as discussed above and relayed to the host computer as
difference information to augment or provide cursor and/or scene movement in
20 response to the user's movement.

Note that such edge movements roughly correspond to movement along
rows and columns of the detector array. Detected movement from one row to
another (or one column to another) can readily calculate the actual motion of the
25 user from information of the user's best focus position and from the focal length of
the camera's lens. This information may then be used to set the magnification of
movement of items in the computer display (e.g., cursor and/or scene view).

Figure 16 illustrates an image of one object 500 used in accord with the
30 invention to provide image manipulation in response to motion of the object. The
object 500 is held by the user 501 to manipulate motion of his cursor 502 and/or

scene view 504 on his computer display 506. The object 500 is used because it exhibits an optical shape that is easily recognized through image correlation (such as matched filtering). In accord with the invention, a camera 510 is used to image the object 500; and frames of data are sent to the frame processor 512. The processor 512
5 determines image position - relative to a starting position - and thereafter communicates difference information to the user's computer 505 along data line 514. The difference information is used by the computer's CPU and operating system to reposition items on the display 506 in response to motion of the object 500. Almost any motion, including rotation, tilting and translation are accomplished with the
10 object 500 relative to a start position. This start position can be triggered by the user 501 at the start of a game by commanding that the camera 510 take a reference frame ("REF") that is stored in memory 513. The user 501 commands that REF imagery be taken and stored through the keyboard 505a, connected to the computer 505, which in turn commands the processor 512 and camera 510 to take the reference frame REF.

15

Motion of the object 500 is thus made possible with enhanced accuracy by comparing subsequent frames of the object 500 with REF. When motion of rotation, tilt or translation are detected (for example, by using the techniques of Figures 2-4, 8-8D), then repositioning of items (502, 504) on the display 506 are follow that
20 movement.

The techniques of the invention permit control of the scene view and/or cursor on a computer screen by motion of one or more parts of the user's body. Accordingly, as shown in Figure 17, complete motion of the user 598 can be
25 replicated, in the invention, by correlated motion of an action figure 599 within a game. In Figure 17, user 598 is imaged by a camera 602 of the invention; and frames from the camera 602 are processed by process section 604, such as described herein. The user 598 is captured and processed, in digital imagery, and annotated with appropriate user segments, e.g., segments 1-6 indicating the user's hands, feet, head
30 and main body. Motion of the segments 1-6 are communicated to the host computer 606 from the process section 604. The computer's operating system then updates the

associated display 608 so that the action figure 599 (corresponding to an action figure within a computer game) moves like user 598. Accordingly, user motion of action figure 599 is made by the user 598 by performing stunts (e.g., striking and kicking) that he would like the action figure 599 to perform, such as to knock out an
5 opponent within the display 608.

In an alternative embodiment, icons can be used to simplify image and motion recognition of user segments such as segments 1-6. For example, if user 598 is marked with a star-shaped object on her hand (e.g., segment 1), then that star symbol
10 is more easily recognized by algorithms such as described herein to determine motion. By way of another example, the hand of user 598 can be covered with a glove that has an "+" symbol on the glove. That "+" symbol can be used to more easily interpret user motion as compared to, for example, actually interpreting motion of the user's hand, which is rounded with five fingers. In a third example,
15 user 598 can wear a article of clothing such as shirt 598a with a "+" symbol 598b; and the invention can be used to track the icon "+" 598b with great precision since it is a relatively easy object to track as compared to actual body parts. It should be apparent to those in the art that icons such as symbol 598b can be painted or pasted onto the individual too to obtain similar results.

20

Figure 18 illustrates a two camera system 700 used to determine translation and rotation. The forward viewing camera 702 observes the user's face 703 and determines the right-left (Δx_1) and up-down (Δy_1) translation of the user's face 703. The top viewing camera 704 observes the top of the user's face or head 705 and
25 determines the right-left (Δx_2) and forward-backward motion (Δy_2) of the user's face or head. The two cameras 702, 704 are each processed through motion sensing algorithms 706 using the teachings above, and results are shown on the computer display 710. For purposes of illustration, the display 710 shows an image of the user; while the image can be, for example, an action figure or other computer object
30 (including the computer cursor), as desired, which follows tracking motions Δx_1 , Δy_1 , Δx_2 , Δy_2 . As indicated in Figure 18, for example, Δy_2 can be directly applied to

motion control of the user's forward and reverse motion (note, these motions are illustrated as within a computer display 710 as processed by algorithms 706). Δx_1 can be directly applied to the users left-right sideways or strafe motion; Δy_1 can be directly applied to control the users up-down viewpoint, each as illustrated on display 710a. The results of the difference between Δx_2 and Δx_1 can be applied to control the user's left-right turn or viewpoint.

The techniques of Figure 18 can be further extended to front, side and top view cameras for complete motion detection. The top camera determines the user's left-right, front-back motion while the front facing camera determines the user's rotational up-down, left-right motion.

Figure 19 describes an algorithm to detect user eye blink. The video imagery is stored into a multiple frame buffer 800. The algorithm selects the current frame and a frame from the frame buffer and differences these frames using the adder 802. The difference frame consists of the pixel by pixel difference of the delayed frame and the current frame. The difference frame includes motion information used by the algorithms of teachings above. It also contains information on the user eye blink. The frames differenced by the adder 802 are separated temporally enough to ensure that one frame contains an image of the users face with the eyes open, the other image is of the user's face with the eyes closed. The difference image contains a two strong features, one for each eye. These features are spatially separated by the distance between the user's eyes. The blink detect function 808 inspects the image for this pair of strong features which are aligned horizontally and spaced within an expected distance based on the variation from one human face to another and the variation in seating distance expected from user to user. The recognition of the blink features may be accomplished using a matched filter or by recognition of expected frequency peaks in the frequency domain at the expected spatial frequency for human eye separation. The blink detect function 708 identifies the occurrence of a blink to a controlling function to either disable the cursor motion or take some other action.

Figure 20 illustrates a sound re-calibration system 800 constructed according to the invention. As above, a camera 802 is arranged to view a user, a part of a user (e.g., a hand), or an object through the camera's field of view 804. A processing section 806 correlates the framing image data from camera 802 to induce movement of a scene view or cursor on the user's display 810. For purposes of illustration, the scene view or cursor is shown illustratively as a dot 808 on display 810; and movement 812 of the cursor 808 from position 808a to 808b represents a typical movement of the cursor or scene view 808 in response to movement within the field of view 804, as described above. A re-calibration section 816 is used to reset the cursor or scene view 808 back to an initial position 808a, if desired. Specifically, in one embodiment, section 816 is a microphone that responds to sound 818 generated from a sound event 818a, such as a snap of the user's fingers, or a particular word uttered by the user, to generate a signal for processing section 806 along signal line 816a; and section 806 processes the signal to move the cursor or scene view 808 back to original position 808a. In another embodiment, re-calibration section 816 can also correspond to a processing section within the processing hardware/software of system 800 to, for example, respond to the blink of a user's eyes to cause movement of the cursor 808 back to position 808a.

20

The invention thus attains the objects set forth above, among those apparent from the preceding description. Since certain changes may be made in the above methods and systems without departing from the scope of the invention, it is intended that all matter contained in the above description or shown in the accompanying drawing be interpreted as illustrative and not in a limiting sense. By way of example, although FFT correlation is often discussed in the above description, it should be apparent to those skilled in the art that other correlation techniques can be used with the invention to achieve similar results without departing from the scope hereof.

30

The following Matlab source code provides non-limiting computer code

suitable for use to control the cursor on a display such as described herein. The Matlab source code thus provides an operational demonstration of the concepts described and claimed herein. The Matlab source code is platform independent and needs only a sequence of input images. It includes a centroid operation on the correlation peak which is not included on the PC version (described below),
 5 providing a finer measurement on the motion in the image. More particularly, the centroid operation provides a refinement on locating the correlation peak. The PC code, discussed below, uses the pixel location nearest the correlation peak while the centroiding operation improves the resolution of the peak location to levels below a
 10 pixel.

```

%           Copyright (C) 1998
%           Video Mouse Group Partnership
%
15 % This following script file reads in a sequence of images of a computer user's face.
% It then processes the image sequence using the methods of difference
% frame correlation processing used for a human-computer interface.
% This code includes a centroiding operation and demonstrates the
% difference frame correlation approach.
20
[filename,filepath]=uigetfile('d:\videomouse\*.mat');
files=dir([filepath '*.mat']);
previousFrame=zeros(64);
inputMatrix1=zeros(64);
25 inputMatrix2=zeros(64);
cursorPosX(1)=0;
cursorPosY(1)=0;
for i=1:size(files,1);
    load([filepath 'frame' num2str(i)]);
30     camera=conv2(double(camera),ones(2)/4,'same');
    camera=camera(1:2:480,1:2:640);
    camera=round(camera(120:32:120+31,160:32:160+31)/4);
    if mod(i,2)
        %Compute difference frame
35     inputMatrix2=camera-previousFrame;
        %Save current difference frame for next iteration
        previousFrame=camera;
        %FFT difference frame
        inputMatrix2=fft2(inputMatrix2);
40     %Perform difference frame correlation by multiplying difference frame by
        %complex conjugate of previous frame

```



```

correlationMatrix=0.00001*conj(inputMatrix2.*conj(inputMatrix1));
else
    %Compute difference frame
    inputMatrix1=camera-previousFrame;
5    %Save current difference frame for next iteration
    previousFrame=camera;
    %FFT difference frame
    inputMatrix1=fft2(inputMatrix1);
    %Perform difference frame correlation by multiplying difference frame by
10    % complex conjugate of previous frame
    correlationMatrix=0.00001*conj(inputMatrix1.*conj(inputMatrix2));
end
%Compute inverse FFT on correlation matrix
correlationMatrix=real(fft2(correlationMatrix));
15 %Shift correlation matrix to center correlation peak
temp=fftshift(correlationMatrix);
%Find maximum value of correlation matrix
correlationPeak=max(correlationMatrix(:));
%Perform centroiding on correlation peak in order to find peak location
20 %The centroiding operation is not currently incorporated into the application version
mask=temp>.50*correlationPeak(1);
maskSize(i)=sum(mask(:));
if maskSize(i)<100
    colSums=sum((temp-.50*correlationPeak(1)).*mask);
25    xPos(i)=sum(colSums.*(-32:31))/sum(colSums);
    rowSums=sum((temp'-.50*correlationPeak(1)).*mask');
    yPos(i)=sum(rowSums.*(-32:31))/sum(rowSums);
else
    xPos(i)=0;
30    yPos(i)=0;
end
if i>1
    cursorPosX(i)=cursorPosX(i-1)+xPos(i);
    cursorPosY(i)=cursorPosY(i-1)+yPos(i);
35 end
end
end

```

The following PC source code, labeled videoMouseDlg.doc and videoMouseDSP.doc, provide non-limiting and nearly operable DSP code for control of the cursor, as described herein. The code is not smooth; and there are other files required to compile this code to an executable, as will be apparent to those skilled in the art, including header files (*.h), resource files and compiler directives.

```

*
*           Copyright (C) 1998
*           Video Mouse Group Partnership
*
10  *
*   Module : dspcode.c

/**===== MODIFICATION HISTORY =====**/
/*
15 */

/**===== MODULE ENVIRONMENT =====**/
/*
* Include files
20 */

#include "dddefs.h"      /* XPG definitions and prototypes. */
#include "ddptypes.h"    /* XPG function prototypes. */
#include "dberrors.h"    /* XPG error codes. */
25 #include "xpreg.h"     /* XPG register definitions */
#include "grabisr.h"     /* grab isr include file */
#include "intpt40.h"     /* Parallel Runtime Support Library intertuupts header */

#include "protocol.h"    /* Protocol constants defined for this application */
30

/**=====**/
/*
* Function prototypes.
*/
35

extern VOID   DBU_App1Interrupt (VOID);
extern VOID   DDK_PKTDelay (ULONG);

/**=====**/
40 /*
* Global Variables
*/

```

volatile LONG G_IApp1UserMaskIntCount = 0;

/**===== CODE =====**/

/*=====**/

```

5  *
  * Name      : DBU_Correlation
  *
  * Parameters :
  *   InputImage1  Image number of the 1st input image.
10 *   InputImage2  Image number of the 2nd input image.
  *
  * Returns   : Error status
  *             P_SUCCESS
  *
15 * IMotionX
  * IMotionY
  * IFrameRate
  *
  *=====**/
20

```

/* FFT STUFF FROM TI */

```

#define FFTSIZE      64          /* FFT size (n)          */
25 #define LOGFFTSIZE  6          /* log(FFT size)        */
#define FFTSIZEEx2   128
#define BLOCK0       0x002ff800 /* on-chip RAM buffer   */

extern void cfftcc();          /* C-callable complex FFT */

30 float complexMatrix1[FFTSIZE][FFTSIZEEx2]; /* Input matrix */
float complexMatrix2[FFTSIZE][FFTSIZEEx2]; /* Input matrix */
float correlationMatrix[FFTSIZE][FFTSIZEEx2];
long previousFrame[FFTSIZE][FFTSIZE];
35 float *p_localRam,*fPtr1,*fPtr2;
float correlationPeak;
float *block0 = (float *)BLOCK0, *mm1[FFTSIZE], *mm2[FFTSIZE], *mm3[FFTSIZE];

```

```

40
/*End FFT Stuff*/

```

```

LONG DBU_UserFunction (LONG lInputImage1, LONG lInputImage2)
{

```

```

45     PROCESSINGINFO ImageInfo1;
     PROCESSINGINFO ImageInfo2;

```

```

    LONG  IErrorStatus;
    LONG  IFifoStatus;
    LONG  IOutputFifoStatus;
5    LONG  IFrameCount = 0;
    LONG  IStatus;
    register LONG  *pIAddress1;
    register LONG  *pIAddress2;
    register LONG  *p_imageAddress;
10    long *IPtr;
    float *currentDifference;
    float *previousDifference;

    ULONG  ulPCData;
15    ULONG  ulTemp0;
    ULONG  ulValue;
    ULONG  pulPacket[5];

    long peakRow;
20    long peakCol;
    long row;
    long col;
    long pixel;
    long index1;
25    long index2;
    long index3;

    int targetCol= 16,targetRow = 16,targetDirection=0;
30

    /*-----*/
    /* Initialize some variables. */
    /*-----*/

35    IErrorStatus = P_SUCCESS;
    ulPCData = APPLICATION_RUNNING;
    IFifoStatus = P_EMPTY;

    /*-----*/
40    /* Install the ISR found in the file INTERRUPT.ASM to interrupt */
    /* resource IIOF0. Initialize the G_IApp1UserMaskIntCount. */
    /*-----*/

    DDF_ISRSetIIOF0 (P_INTERRUPT_USER_MASK, (VOID *)
45    DBU_App1Interrupt);
    G_IApp1UserMaskIntCount = 0;

```

```

/*-----*/
/* Set the ID of the second image for double buffering.      */
/* Perform a quick grab setup on the input image number.    */
/* Call Quick Grab with no wait. This application runs the grab */
5 /* in continuous mode, the grab will not return until the DMA */
/* has started.                                             */
/*-----*/

    if ((IStatus = DBF_SetSecondImageID (P_DEFAULT_QGS, IInputImage2))
10      != P_SUCCESS)
    {
        ulPCData = END_APPLICATION_REQUEST;
        IErrorStatus = IStatus;
    } /* End if. */

15
    if ((IStatus = DBF_QuickGrabSetup (P_DEFAULT_QGS, IInputImage1))
        != P_SUCCESS)
    {
        ulPCData = END_APPLICATION_REQUEST;
20        IErrorStatus = IStatus;
    } /* End if. */

    if ((IStatus = DBF_QuickGrab (P_DEFAULT_QGS, P_GRAB_INIT,
25      P_GRAB_NO_WAIT))
        != P_SUCCESS)
    {
        ulPCData = END_APPLICATION_REQUEST;
        IErrorStatus = IStatus;
    } /* End if. */

30
/*-----*/
/* Enable hardware interrupts on IIOF0 */
/*-----*/

35
    INT_ENABLE ();
    set_iif_flag (IIOF0_EIIOF);

/*-----*/
/* Initialize pointers to the two image buffers */
40 /*-----*/

    if ((IStatus = DBK_MintGetImageInfo (IInputImage1, &ImageInfo1))
        != P_SUCCESS)
    {
45        ulPCData = END_APPLICATION_REQUEST;
        IErrorStatus = IStatus;
    } /* End if. */

```

```

        if ((IStatus = DBK_MmtGetImageInfo (lInputImage2, &ImageInfo2))
            != P_SUCCESS)
        {
5           ulPCData = END_APPLICATION_REQUEST;
           IErrorStatus = IStatus;
        } /* End if. */

        plAddress1 = (LONG *) (ImageInfo1.PRO_MappedAddress);
10       plAddress2 = (LONG *) (ImageInfo2.PRO_MappedAddress);

        /* FFT Initialization Stuff */
        asm(" or 1800h,st");          /* cache enable          */
        /*End FFT Initialization */
15

        /*-----*/
        /* While the input fifo from the PC does not contain */
        /* any data, continue processing frames and returning */
20       /* the results to the PC.          */
        /*-----*/

        while (ulPCData != END_APPLICATION_REQUEST){
25           if (lFrameCount < (G_lApp1UserMaskIntCount)){
               lFrameCount = G_lApp1UserMaskIntCount;

               p_imageAddress = (G_lApp1UserMaskIntCount % 2)
               ? plAddress1 : plAddress2;
30           if (G_lApp1UserMaskIntCount % 2){
               p_imageAddress=plAddress2; /*
               currentDifference=&complexMatrix2[0][0];
               previousDifference=&complexMatrix1[0][0];
           }
           else{
35               p_imageAddress=plAddress1; /*
               currentDifference=&complexMatrix1[0][0];
               previousDifference=&complexMatrix2[0][0];
           }
40

        /****** Compute FFT on Difference Frame Rows *****/

        for (row=0;row<FFTSIZE;row++) {
            for (col=0;col<FFTSIZE;col++){
45                lPtr=p_imageAddress+2*col+256*row;
                pixel=*lPtr+*(lPtr+1)+*(lPtr+128)+*(lPtr+129);

```

```

/*Compute Difference Frame */
block0[2*col]=(float) (pixel -
previousFrame[row][col]);

5      block0[2*col+1]=0.0;
/*Save current frame for next iteration*/
previousFrame[row][col]=pixel;
    }
    cfft(block0,FFTSIZE,LOGFFTSIZE);
    fPtr1=currentDifference+row*FFTSIZE*2;
10    for (index1=0;index1<FFTSIZE*2;index1++)
        fPtr1[index1]=block0[index1];
    }

15    for (col=0;col<FFTSIZE;col++) {

        index3=2*col;
        for (index2=0;index2<FFTSIZE*2;index2=index2+2){
            block0[index2]=currentDifference[index3];
20            block0[index2+1]=currentDifference[index3+1];
            index3+=FFTSIZE*2;
        }
        /*Complete column FFT of difference frame*/
        cfft(block0,FFTSIZE,LOGFFTSIZE);
25

        index3=2*col;
        for(index2=0;index2<FFTSIZE*2;index2=index2+2){

30            /*Save FFT of difference frame */
            currentDifference[index3]=block0[index2];
            currentDifference[index3+1]=block0[index2+1];

            block0[index2]=currentDifference[index3]
35            *previousDifference[index1]
            +currentDifference[index3+1]
            *previousDifference[index1+1];
            block0[index2+1]=currentDifference[index3]
            *previousDifference[index1+1]
40            -currentDifference[index3+1]
            *previousDifference[index1];
            index3+=FFTSIZE*2;
        }
        /*Compute inverse FFT for correlation matrix column*/
45        cfft(block0,FFTSIZE,LOGFFTSIZE);

        /*Save to correlation frame*/

```

```

        fPtr1=&correlationMatrix[0][0];
        index3=2*col;
        for (index2=0;index2<FFTSIZEx2;index2=index2+2){
5             fPtr1[index3]=block0[index2];
              fPtr1[index3+1]=block0[index2+1];
              index3+=FFTSIZEx2;
        }
    }
10    correlationPeak=-100000000;
    for (row=0;row<FFTSIZE;row++) {
        fPtr1=&correlationMatrix[row][0];
        for (index1=0;index1<FFTSIZEx2;index1++)
            block0[index1]=fPtr1[index1];
15    /* Inverse FFT on correlation matrix row */
        cfft(block0,FFTSIZE,LOGFFTSIZE);

        fPtr1=&correlationMatrix[row][0];
        for (col=0;col<FFTSIZE;col++){
20            index1=2*col;
            if (correlationPeak<block0[index1]){
                correlationPeak=block0[index1];
                peakRow=row;
                peakCol=index1;
25            }
            fPtr1[index1]=block0[index1];
        }
    }

30    /*-----*/
    /* Send protocol, lAverage and lFrameCount to the PC. */
    /*-----*/

35    pulPacket[0] = APPLICATION_RUNNING;
    pulPacket[1] = peakCol;
    pulPacket[2] = peakRow;
    pulPacket[3] = (long) (correlationPeak*.0001);
    DDK_PKTSend (P_PACKET_USER_INTERFACE,
40    pulPacket,
        4 * sizeof (LONG),P_WAITFORCOMPLETE,
        P_PC_INTERRUPT);
    } /* End if. */

45    DDK_PKTInterfaceStatus (P_PACKET_USER_INTERFACE, &lFifoStatus,
        &lOutputFifoStatus);
    if (lFifoStatus != P_EMPTY){

```



```

                                DDK_PKTRecv (P_PACKET_USER_INTERFACE,
pulPacket,
                                4 * sizeof (LONG),P_WAITFORCOMPLETE,
                                P_NO_PC_INTERRUPT);
5                                ulPCData = pulPacket[0];
                                } /* End if. */
                                } /* End while. */

                                /*-----*/
10                                /* Disable the IIOF0 interrupt. */
                                /*-----*/

                                reset_iif_flag (IIOF0_EIIOF);
                                DDF_ISRDisableIIOF0 ();

15                                /*-----*/
                                /* Abort the continuous grab. */
                                /*-----*/

20                                DBF_AbortGrab ();
                                DBF_QuickGrabStatus (P_GRAB_WAIT);

                                /*-----*/
                                /* Send back a protocol word indicating the */
25                                /* last packet of data. (Pad to correct size) */
                                /*-----*/

                                ulValue = APPLICATION_TERMINATED;
                                DDK_PKTSend (P_PACKET_USER_INTERFACE, &ulValue, 1L,
30                                P_WAITFORCOMPLETE,
                                    P_NO_PC_INTERRUPT);
                                DDK_PKTSend (P_PACKET_USER_INTERFACE, &IErrorStatus, 1L,
                                    P_WAITFORCOMPLETE, P_NO_PC_INTERRUPT);
                                DDK_PKTSend (P_PACKET_USER_INTERFACE, &IErrorStatus, 1L,
35                                P_WAITFORCOMPLETE,P_PC_INTERRUPT);

                                /*-----*/
                                /* Empty USER input FIFO and output FIFO. */
                                /* (Host won't get the END message until */
40                                /* the output FIFO is empty !) */
                                /*-----*/

                                DDK_PKTInterfaceStatus (P_PACKET_USER_INTERFACE, &IFifoStatus,
                                    &IOutputFifoStatus);

45                                while ((IFifoStatus != P_EMPTY) || (IOutputFifoStatus != P_EMPTY)){
                                    if (IFifoStatus != P_EMPTY){

```

```

        DDK_PKTRecv (P_PACKET_USER_INTERFACE,
&ulPCData, 1L,
        P_WAITFORCOMPLETE, P_NO_PC_INTERRUPT);
    } /* End if. */
5    DDK_PKTInterfaceStatus (P_PACKET_USER_INTERFACE, &IfifoStatus,
        &lOutputFifoStatus);
    } /* End while. */

    return P_SUCCESS;

10 } /* End of the DBU_UserFunction function. */

/**=====**/

15 /*
 *      Copyright (C) 1998
 *      Video Mouse Group Partnership
 */

20 /**=====
 **
 */
// videomouseDlg.cpp : implementation file
//

25 #include "stdafx.h"
#include "videomouse.h"
#include "videomouseDlg.h"

30 #include "dpdefs.h"    /* XPG definitions and prototypes. */
#include "dpptypes.h"    /* XPG function prototypes. */
#include "dberrors.h"    /* XPG error codes. */
#include "protocol.h"    /* Protocol constants define for this application */
#include <conio.h>        /* getch */
35 #include <math.h>

static int s_runDSPLoopThreadProc;

#ifdef _DEBUG
40 #define new DEBUG_NEW
#undef THIS_FILE
static char THIS_FILE[] = __FILE__;
#endif

45 #define P_ID_USER_FUNCTION1    0L
#define P_PCOUNT_USER_FUNCTION1    2

```

```
#define FRAMESIZE 64
```

```
UINT DSPLoopProc(LPVOID pclass)
{
```

```

5      CVideomouseDlg& pcdd = *(reinterpret_cast<CVideomouseDlg*>(pclass));
      CString dataString;

      int savedCommandMode = DPK_XCCSetCommandType (P_USER);
10     DPK_XCCSetWaitMode (P_NO_WAIT);

      DPK_XCCPushOpcode (P_ID_USER_FUNCTION1,
P_PCOUNT_USER_FUNCTION1);
      DPK_XCCPushLong ((unsigned long) pcdd.m_inputImageNumber2);      /* 2 */
15     DPK_XCCPushLong ((unsigned long) pcdd.m_inputImageNumber1);      /* 1 */
      DPK_XCCSetCommandType (savedCommandMode);
      long status = DPK_XCCCheckStatus (P_PCOUNT, P_XCCFIX);
      double framesPerSecond;
      time_t start, finish;
20     time( &start );
      int counter=0;
      if (status==P_SUCCESS){

          while(s_runDSPLoopThreadProc){
25             counter++;
             status = DPK_PKTRcv (P_PACKET_USER_INTERFACE, (HVOID
*)

                pcdd.m_DSPPacket, 4 * sizeof(LONG),
                P_WAIT_COMPLETE);
30             long protocol = pcdd.m_DSPPacket[0];
                dataString.Format("%d",pcdd.m_DSPPacket[1]);
                pcdd.m_average.SetWindowText(dataString);
                dataString.Format("%d",pcdd.m_DSPPacket[2]);
                pcdd.m_frameNumber.SetWindowText(dataString);
35             time( &finish );
                double elapsedTime = difftime( finish, start );
                framesPerSecond=(double) counter/ (double)elapsedTime;
                long max = pcdd.m_DSPPacket[3];

40             if (max>0.0){
                double detectx=(double) pcdd.m_DSPPacket[2];
                double detecty=0.5*(double)pcdd.m_DSPPacket[1];

                if (detectx > 31) detectx =detectx-FRAMESIZE;
45             if (detecty > 31) detecty = FRAMESIZE-detecty;
                else detecty = -detecty;
            }
        }
    }
}

```

```

//          double multiplier;/
//          if (detectx<2) multiplier=1.0;
//          else if (detectx<10) multiplier = exp((detectx-2.0)/2.5);
//          else multiplier=0;
5 //
//          detectx*=multiplier;
//          if (detecty<2) multiplier=1.0;
//          else if (detecty<10) multiplier = exp((detecty-2.0)/2.5);
//          else multiplier=0;
10 //          detecty*=multiplier;

          static POINT ptCursor;
          GetCursorPos(&ptCursor);

15          ptCursor.x-=(long) detectx;
          ptCursor.y-=(long) detecty;

          SetCursorPos(ptCursor.x,ptCursor.y);
      }
20  }
      savedCommandMode = DPK_XCCSetCommandType
(savedCommandMode);
      DPK_XCCSetWaitMode (P_WAIT_COMPLETEE);
      DPK_EndPCK ();
25      AfxMessageBox("Exited Thread");
    }
    else{
      s_runDSPLoopThreadProc=false;
      savedCommandMode = DPK_XCCSetCommandType
30 (savedCommandMode);
      DPK_XCCSetWaitMode (P_WAIT_COMPLETEE);
      DPK_EndPCK ();
      AfxMessageBox("Exited Thread");
    }
35    time( &finish );
    double elapsedTime = difftime( finish, start );
    framesPerSecond=(double) counter/ (double)elapsedTime;
    return I;
  }
40  //////////////////////////////////////
  // CAboutDlg dialog used for App About

  class CAboutDlg : public CDialog
45  {
  public:
    CAboutDlg();

```

```

// Dialog Data
    //{AFX_DATA(CAboutDlg)
    enum { IDD = IDD_ABOUTBOX };
5    //{AFX_DATA

    // ClassWizard generated virtual function overrides
    //{AFX_VIRTUAL(CAboutDlg)
    protected:
10    virtual void DoDataExchange(CDataExchange* pDX); // DDX/DDV support
    //{AFX_VIRTUAL

// Implementation
protected:
15    //{AFX_MSG(CAboutDlg)
    //{AFX_MSG
    DECLARE_MESSAGE_MAP()
};

20 CAboutDlg::CAboutDlg() : CDialog(CAboutDlg::IDD)
{
    //{AFX_DATA_INIT(CAboutDlg)
    //{AFX_DATA_INIT
25
void CAboutDlg::DoDataExchange(CDataExchange* pDX)
{
    CDialog::DoDataExchange(pDX);
    //{AFX_DATA_MAP(CAboutDlg)
30    //{AFX_DATA_MAP
}

BEGIN_MESSAGE_MAP(CAboutDlg, CDialog)
    //{AFX_MSG_MAP(CAboutDlg)
35    // No message handlers
    //{AFX_MSG_MAP
END_MESSAGE_MAP()

////////////////////////////////////
40 // CVideomouseDlg dialog

CVideomouseDlg::CVideomouseDlg(CWnd* pParent /*=NULL*/)
: CDialog(CVideomouseDlg::IDD, pParent)
{
45    //{AFX_DATA_INIT(CVideomouseDlg)
    //{AFX_DATA_INIT
    // Note that LoadIcon does not require a subsequent DestroyIcon in Win32

```

```

        m_hIcon = AfxGetApp()->LoadIcon(IDR_MAINFRAME);
    }

void CVideomouseDlg::DoDataExchange(CDataExchange* pDX)
{
    CDialog::DoDataExchange(pDX);
    //{{AFX_DATA_MAP(CVideomouseDlg)
    DDX_Control(pDX, IDC_FRAMENUMBER, m_frameNumber);
    DDX_Control(pDX, IDC_AVERAGE, m_average);
    //}}AFX_DATA_MAP
}

BEGIN_MESSAGE_MAP(CVideomouseDlg, CDialog)
    //{{AFX_MSG_MAP(CVideomouseDlg)
    ON_WM_SYSCOMMAND()
    ON_WM_PAINT()
    ON_WM_QUERYDRAGICON()
    ON_BN_CLICKED(IDC_ENABLE, OnEnable)
    ON_BN_CLICKED(IDC_STOP, OnStop)
    //}}AFX_MSG_MAP
END_MESSAGE_MAP()

////////////////////////////////////
// CVideomouseDlg message handlers

BOOL CVideomouseDlg::OnInitDialog()
{
    CDialog::OnInitDialog();

    // Add "About..." menu item to system menu.

    // IDM_ABOUTBOX must be in the system command range.
    ASSERT((IDM_ABOUTBOX & 0xFFF0) == IDM_ABOUTBOX);
    ASSERT(IDM_ABOUTBOX < 0xF000);

    CMenu* pSysMenu = GetSystemMenu(FALSE);
    if (pSysMenu != NULL)
    {
        CString strAboutMenu;
        strAboutMenu.LoadString(IDS_ABOUTBOX);
        if (!strAboutMenu.IsEmpty())
        {
            pSysMenu->AppendMenu(MF_SEPARATOR);
            pSysMenu->AppendMenu(MF_STRING, IDM_ABOUTBOX,
45   strAboutMenu);
        }
    }
}

```

```

        // Set the icon for this dialog. The framework does this automatically
        // when the application's main window is not a dialog
        SetIcon(m_hIcon, TRUE);           // Set big icon
5      SetIcon(m_hIcon, FALSE);          // Set small icon

        InitializeFrameGrabber();

        return TRUE; // return TRUE unless you set the focus to a control
10    }

void CVideoMouseDlg::OnSysCommand(UINT nID, LPARAM lParam)
{
    if ((nID & 0xFFF0) == IDM_ABOUTBOX)
15    {
        CAboutDlg dlgAbout;
        dlgAbout.DoModal();
    }
    else
20    {
        CDialog::OnSysCommand(nID, lParam);
    }
}

25 // If you add a minimize button to your dialog, you will need the code below
// to draw the icon. For MFC applications using the document/view model,
// this is automatically done for you by the framework.

void CVideoMouseDlg::OnPaint()
30 {
    if (IsIconic())
    {
        CPaintDC dc(this); // device context for painting

35        SendMessage(WM_ICONERASEBKGND, (LPARAM) dc.GetSafeHdc(), 0);

        // Center icon in client rectangle
        int cxIcon = GetSystemMetrics(SM_CXICON);
        int cyIcon = GetSystemMetrics(SM_CYICON);
40        CRect rect;
        GetClientRect(&rect);
        int x = (rect.Width() - cxIcon + 1) / 2;
        int y = (rect.Height() - cyIcon + 1) / 2;

45        // Draw the icon
        dc.DrawIcon(x, y, m_hIcon);
    }
}

```

```

        else
        {
            CDialog::OnPaint();
        }
5   }

// The system calls this to obtain the cursor to display while the user drags
// the minimized window.
HCURSOR CVideomouseDlg::OnQueryDragIcon()
10  {
        return (HCURSOR) m_hIcon;
    }

void CVideomouseDlg::OnEnable()
15  {
        s_runDSPLoopThreadProc = true;
        m_pDSPLoopThread = AfxBeginThread (DSPLoopProc, this);
    }

20

void CVideomouseDlg::OnStop()
{
    if(s_runDSPLoopThreadProc){
25        m_DSPPacket[0] = END_APPLICATION_REQUEST;
        DPK_PKTSend (P_PACKET_USER_INTERFACE, m_DSPPacket, 4 * sizeof
            (LONG),
            P_WAIT_COMPLETE);
        s_runDSPLoopThreadProc=false;
30    }
}

#define INIT_FAILURE 1
long CVideomouseDlg::InitializeFrameGrabber()
{
35    DPK_InitPCK(1);

    if ((m_status=DPK_InitXPG (0, P_IFB_RELOAD_COFF_FILE |
        P_IFB_CHECK_REVISION,"videoMouse.out")) != P_SUCCESS){
        DPK_EndPCK ();
40        m_errorMessage.Format("Error initializing FPG, status = %ld.\n", m_status);
        AfxMessageBox(m_errorMessage);
        return INIT_FAILURE;
    }
    DPK_XCCSetWaitMode (P_WAIT_COMPLETE);
45    long cpsNumber = DPF_LoadCPF("vidmouse.cpf");
    if (cpsNumber != P_SUCCESS){
        m_errorMessage.Format("Error loading a CPF, status = %ld.\n", m_status);
    }
}

```



```

        AfxMessageBox(m_errorMessage);
        DPK_EndPCK ();
        return INIT_FAILURE;
    }
5    if ((m_status = DBF_SelectCPS (cpsNumber)) < P_SUCCESS){
        m_errorMessage.Format("Error selecting a CPS, status = %ld.\n", m_status);
        AfxMessageBox(m_errorMessage);
        DPK_EndPCK ();
        return INIT_FAILURE;
10    }

    m_status=DBF_SetGrabWindow(P_DEFAULT_QGS, 256,128,176,128);

    if ((m_status = DBF_GetGrabWindow (P_DEFAULT_QGS, &m_startCol,
15    &m_numCols,
        &m_startRow, &m_numRows)) != P_SUCCESS){
        m_errorMessage.Format("Error DBF_GetGrabWindow: %ld.\n", m_status);
        AfxMessageBox(m_errorMessage);
        DPK_EndPCK ();
20    return INIT_FAILURE;
    }

    if ((m_status = DBK_MmtCreateImage (m_numCols, m_numRows,
P_DATA_SIZE_BYTE,
25    P_DATA_TYPE_INTEGER, 2, &m_inputImageNumber1,
    &m_numberImagesCreated))
        != P_SUCCESS){
        m_errorMessage.Format("Error creating the input image; status = %ld.\n",
m_status);
30    AfxMessageBox(m_errorMessage);
        DPK_EndPCK ();
        return INIT_FAILURE;
    }
    m_inputImageNumber2 = m_inputImageNumber1 + 1;
35    return 0;
}

```

40 It should be understood that the following claims are to cover all generic and specific features of the invention described herein, and all statements of the scope of the invention which, as a matter of language, might be said to fall there between.

Having described the invention, what is claimed is:

1. A human motion following controller for augmenting motion of items shown on a computer display, the display being coupled to a computer of the type which controls positioning of the items through operating system controls, comprising:
 - 5 a camera for capturing frames of data corresponding to a first image of at least part of a user at the computer display;

signal processing means coupled to the camera for (a) detecting differences between successive frames of data corresponding to motion of the first image, and (b)

 - 10 communicating differences information to the computer to reposition display of the items through operating system controls, the items being repositioned on the display by an amount corresponding to the motion of first image.
2. A controller of claim 1, wherein the items comprise a computer cursor.
 - 15 3. A controller of claim 1, wherein the items comprise a scene view.
 4. A controller of claim 1, further comprising a PC card for installation within the computer and for communication on a computer bus, the signal processing
 - 20 means being substantially resident with the PC card for communicating differences information to the bus.
 5. A controller of claim 1, wherein the camera comprises means for capturing augmented frames of data corresponding to a second image of part of the user at the
 - 25 computer display, the signal processing means further comprising means for detecting differences between successive augmented frames of data corresponding to motion of the second image and for communicating augmented difference information to the computer to reposition display of the items through operating system controls, the items being repositioned on the display by an amount
 - 30 corresponding to motion of the first and second images.

6. A controller of claim 1, further comprising frame difference electronics for storing and subsequently subtracting pixel-by-pixel difference data.
7. A controller of claim 6, wherein the difference electronics comprise multiple
5 frame memory, a subtraction circuit, and a state machine controller/memory addresser to control data flow.
8. A controller of claim 1, further comprising N frame video memory for storing frames of image data.
- 10
9. A controller of claim 1, further comprising a DSP for implementing select algorithms on difference frames or raw frames of image data.
10. A controller of claim 9, further comprising memory selected from the group
15 of EPROM and RAM.
11. A controller of claim 9, further comprising means for interfacing the DSP to a PCI bus in the computer.
- 20 12. A controller of claim 1, further comprising MPEG compression electronics for compressing video for the computer.
13. A controller of claim 1, wherein the signal processing means comprises frame differencing means for removing unchanged information from image frames.
- 25
14. A controller of claim 1, wherein the signal processing means comprises frame memory to buffer one or more image frames.
15. A controller of claim 14, wherein the signal processing means comprises a
30 frame differencer for reading a delayed frame from the frame memory and for subtracting the delayed frame from a current image frame.

16. A controller of claim 1, wherein the signal processing means comprises correlation means for determining row and column shifts corresponding to differences between a current image frame and a delayed image frame.

5 17. A controller of claim 16, wherein the signal processing means comprises best fit algorithm means for minimizing the shifts to provide alignment.

18. A controller of claim 17, wherein the best fit algorithm means utilizes a peak detect algorithm.

10

19. A controller of claim 1, wherein the signal processing means comprises video cursor control for enabling and alternatively disabling cursor control.

20. A controller of claim 19, wherein the video cursor control comprises means
15 responsive to keystrokes at the computer.

21. A controller of claim 19, wherein the video cursor control comprises means responsive to a blink of an eye of the user.

20 22. A controller of claim 19, wherein the video cursor control comprises means responsive to sound generated by the user.

23. A controller of claim 22, further comprising a microphone to detect the sound.

25 24. A controller of claim 1, wherein the signal processing means comprises a complex multiplier for providing a two dimensional inverse FFT operation.

25. A controller of claim 24, wherein the signal processing means comprises a peak detect for determining a shift associated with aligning difference images.

30

26. A controller of claim 1, wherein the signal processing means comprises FFT means for providing a two dimensional FFT of image data.
27. A controller of claim 1, wherein the signal processing means comprises means
5 for identifying parts of the user, the parts being selected from the group of a hand, elbow, head, neck, ears, and forehead.
28. A controller of claim 1, wherein the signal processing means comprises means
10 for detecting left and right movement of a head of the user and for shifting the items in response to the left and right movement.
29. A controller of claim 1, wherein the signal processing means comprises means
15 for detecting rotational movement of a head of the user and for rotating the items in response to the left and right movement.
29. A controller of claim 1, wherein the signal processing means comprises means
for repositioning the items, if appropriate, at approximately every 1/30th of a second.
30. A controller of claim 1, wherein the signal processing means comprises means
20 for repositioning the items at a selected magnification as compared to actual movement of the user.
31. A controller of claim 1, wherein the signal processing means comprises means
25 for storing image data of a head of the user at various orientations relative to the camera and for correlating image data to the stored image data to define head orientation, the head orientation being used to reposition the items.
32. A controller of claim 1, wherein the signal processing means comprises IR
30 means for detecting heat associated with the user and for repositioning the items at a rate correlated to the heat.

33. A controller of claim 32, wherein the items correspond to computer gaming display images.
34. A controller of claim 32, wherein the heat corresponds to user stress.
- 5 35. A controller of claim 1, further comprising at least one other camera arranged to take images of at least a second part of the user.
36. A controller of claim 35, wherein the one other camera takes image data in a
10 second electromagnetic spectrum.
37. A controller of claim 1, wherein the camera comprises a DSP.
38. A controller of claim 37, wherein the DSP processes difference information for
15 the computer.
39. A controller of claim 1, wherein the signal processing means comprises a CPU within the computer.
- 20 40. A controller of claim 1, wherein the signal processing means comprises means for processing multiple image zones in frames of image data and for repositioning the items according to characteristics between zones.
41. A controller of claim 40, wherein one zone comprises image data
25 corresponding to at least one eye of the user.
42. A controller of claim 41, wherein the signal processing means comprises means for determining a blink of the eye.
- 30 43. A controller of claim 42, wherein the signal processing means comprises means for disabling and alternatively enabling cursor control based upon the blink.

44. A controller of claim 1, wherein the signal processing means comprises means for processing image data to determine motion of at least one eye of the user and for repositioning the items based upon the motion.
- 5 45. A controller of claim 1, wherein the signal processing means comprises means for processing image data to determine motion of a pupil of at least one eye of the user and for repositioning the items based upon the motion.
46. A controller of claim 1, wherein the camera comprises a zoom attachment for
10 automatically zooming into a desired magnification of at least one eye of the user.
47. A controller of claim 1, wherein the camera comprises zoom means for automatically focusing on the user as the user moves in distance from the camera.
- 15 48. A controller of claim 47, wherein the signal processing means comprises means for enlarging or shrinking the items on the display in response to focusing by the zoom means.
49. A controller of claim 1, wherein the signal processing means comprises means
20 for determining edges of a head of the user and for repositioning the items in response to movements of the edges.
50. A controller of claim 1, wherein the signal processing means comprises means for isolating one or more objects held by the user and for repositioning the items in
25 response to movement of the objects.
51. A controller of claim 1, wherein the signal processing means comprises means for isolating one or more parts of the user and for repositioning the items in response to movement of the parts.

52. A controller of claim 1, wherein the parts comprise at least one of a hand, head, and a foot.

53. A controller of claim 1, wherein the signal processing means comprises means
5 for isolating one or more symbols associated with the user and for repositioning the items in response to movement of the symbols.

54. A controller of claim 1, further comprising a second camera constructed and arranged for viewing the user from above, the signal processing means having
10 means for repositioning the items in response to movement detected from images in the second camera.

55. A controller of claim 54, wherein signal processing means comprises means for repositioning the items in response to forward and backward movement of the
15 user as detected by the second camera.

56. A controller of claim 1, further comprising re-calibration means connected with the signal processing means for repositioning the items to an original position in response to a re-calibration event.

20 57. A controller of claim 1, wherein the re-calibration means comprises a microphone and the event comprises a sound generated by the user.

58. A controller of claim 1, wherein the camera comprises the re-calibration
25 means for detecting a blink of the user.

59. A system for controlling a computer, comprising:

a transducer for converting optical signals to electrical signals;

electronic means for converting electronic signals to digital data;

signal processor means for detecting motion in the digital data and providing a digital representation of said motion;

communication means for entering one or more of the electronic signals, digital data,
5 and digital representation into the computer to manipulate a computer display in response to the motion.

60. A system of claim 59, wherein the computer comprises the signal processor means.

61. A system of claim 59, wherein said signal processor comprises a digital signal
10 processor separate from a CPU within the computer.

62. A system of claim 59, wherein the transducer, electronic means and signal processor are constructed and arranged into a single device in communication with the computer.

63. A system of claim 59, wherein the communication means comprises one of
15 RS170 video, a PCI bus interface, a digital computer interface, a serial computer interface.

64. A system of claim 59, further comprising means for repositioning a computer cursor in response to the motion.

65. A system of claim 59, wherein the transducer comprises one or more of a
20 visible CCD camera and an IR camera.

66. A system of claim 59, wherein the transducer comprises a CCD camera having at least 2x2 imaging pixels.
67. A system of claim 66, wherein the camera comprises optics with various fields of view.
- 5 68. A system of claim 59, wherein the transducer comprises one of a CCD or a CMOS integrated circuit with digital outputs.
69. A system of claim 68, wherein the transducer generates RS170 output.
70. A system of claim 68, wherein the transducer generates RS170 digital output.
71. A system of claim 68, wherein the transducer generates digital resolutions of 4
10 bits or greater
72. A system of claim 59, wherein the signal processor comprises a video frame memory.
73. A system of claim 59, wherein the signal processor comprises frame difference functionality.
- 15 74. A system of claim 59, wherein the signal processor comprises video frame difference memory.
75. A system of claim 59, wherein the signal processor comprises correlation functionality.
76. A system of claim 59, wherein the signal processor comprises means for
20 determining best fit motion.

77. A system of claim 59, further comprising means for controlling cursor movement.

78. A system of claim 59, further comprising means for segmenting video images to provide multiple digital representations of the motion corresponding to different portions of the digital representation.

79. A system of claim 78, wherein the optical signals are generated through image acquisition of a portion of a human.

80. A system of claim 78, wherein the optical signals are generated by viewing multiple features of a human.

81. A system of claim 59, further comprising neural net means for learning user motion over time.

TENT COOPERATION TRE.

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

From the INTERNATIONAL BUREAU

To:

Assistant Commissioner for Patents
United States Patent and Trademark
Office
Box PCT
Washington, D.C. 20231
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 12 August 1999 (12.08.99)	
International application No. PCT/US99/00086	Applicant's or agent's file reference 4521/001PC
International filing date (day/month/year) 04 January 1999 (04.01.99)	Priority date (day/month/year) 06 January 1998 (06.01.98)
Applicant FREY, Robert, D. et al	

1. The designated Office is hereby notified of its election made:

☒ in the demand filed with the International Preliminary Examining Authority on:

20 July 1999 (20.07.99)

☐ in a notice effecting later election filed with the International Bureau on:
2. The election ☒ was
☐ was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

<p>The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland</p> <p>Facsimile No.: (41-22) 740.14.35</p>	<p>Authorized officer Kiwa Mpay</p> <p>Telephone No.: (41-22) 338.83.38</p>
--	---

PATENT COOPERATION TREATY

From the.
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To: CURTIS A. VOCK
DUFT, GRAZIANO & FOREST, P.C.
1790-30TH STREET
SUITE 140
BOULDER CO 80301-1018

PCT

NOTIFICATION OF TRANSMITTAL OF INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Rule 71.1)

Date of Mailing <i>(day/month/year)</i>		24 JUL 2000
Applicant's or agent's file reference 4521/001PC		IMPORTANT NOTIFICATION
International application No. PCT/US99/00086	International filing date <i>(day/month/year)</i> 04 JANUARY 1999	Priority Date <i>(day/month/year)</i> 06 JANUARY 1998
Applicant THE VIDEO MOUSE GROUP		

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.
2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.
3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.
4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Authorized officer

PAUL BELL



Receiving No. (703) 306 3330

Telephone No. (703) 306 3019

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 4521/001PC	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/US99/00086	International filing date (day/month/year) 04 JANUARY 1999	Priority date (day/month/year) 06 JANUARY 1998
International Patent Classification (IPC) or national classification and IPC IPC(7): G09G 5/00 and US Cl.: 345/156		
Applicant THE VIDEO MOUSE GROUP		

- This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
- This REPORT consists of a total of 5 sheets.
☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).
 These annexes consist of a total of 0 sheets.
- This report contains indications relating to the following items:
 - ☒ Basis of the report
 - ☐ Priority
 - ☐ Non-establishment of report with regard to novelty, inventive step or industrial applicability
 - ☐ Lack of unity of invention
 - ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
 - ☐ Certain documents cited
 - ☐ Certain defects in the international application
 - ☐ Certain observations on the international application

Date of submission of the demand 20 JULY 1999	Date of completion of this report 26 JUNE 2000
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer PAUL BELL <i>For [Signature]</i> Telephone No. (703) 306-8019

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US99/00086

I. Basis of the report

1. With regard to the elements of the international application:*

☒ the international application as originally filed☒ the description:

pages 1-58, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of

☒ the claims:

pages 59-68, as originally filed
pages NONE, as amended (together with any statement) under Article 19
pages NONE, filed with the demand
pages NONE, filed with the letter of

☒ the drawings:

pages 1-20, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of

☒ the sequence listing part of the description:

pages NONE, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of

2. With regard to the language, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language _____ which is:☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).☐ the language of publication of the international application (under Rule 48.3(b)).☐ the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

☐ contained in the international application in printed form.☐ filed together with the international application in computer readable form.☐ furnished subsequently to this Authority in written form.☐ furnished subsequently to this Authority in computer readable form.☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.4. ☒ The amendments have resulted in the cancellation of:☒ the description, pages NONE☒ the claims, Nos. NONE☒ the drawings, sheets/fig. NONE5. ☐ This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

* Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US99/00086

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. statement

Novelty (N)

Claims (Please See supplemental sheet) YES
Claims (Please See supplemental sheet) NO

Inventive Step (IS)

Claims (Please See supplemental sheet) YES
Claims (Please See supplemental sheet) NO

Industrial Applicability (IA)

Claims (Please See supplemental sheet) YES
Claims (Please See supplemental sheet) NO

2. citations and explanations (Rule 70.7)

Claims 1-5, 9-11, 14, 30, 35, 37-39, 50-51, 53-55, 59-62, 64, 68, 72, 77-81 lack novelty under PCT Article 33(2) as being anticipated by Platzker et al. (5,528,263).

In regards to claim 1 Platzker shows a human motion following controller for augmenting motion of items shown on a computer display (figure 1, item 14 and 24), the display being coupled to a computer of the type which controls positioning of items through operating system controls (fig. 1, item 12), comprising: a camera for capturing frames of data corresponding to a first image of at least part of a user at the computer display (fig. 1 item 28a), signal processing means coupled to the camera for (fig. 1, item 12b) (a) detecting differences between successive frames of data corresponding to motion of the first image, and (b) communicating differences information to the computer to reposition display of the items through operating system controls, the items being repositioned on the display by an amount corresponding to the motion of first image (column 5, lines 10-32). In regards to claims 2, 64 and 77 Platzker shows wherein the items comprise a computer cursor in response to motion (abstract and figure 5a). In regards to claim 3 Platzker shows wherein the items comprise a scene view (fig. 1, item 24). In regards to claim 4 Platzker does not show the details of a PC card and a computer bus but examiner contends such features are inherent to a computer system as shown in figure 1. In regards to claim 5 Platzker shows the camera comprises means for capturing augmented frames of data corresponding to a second image of a part of the user at the computer display, the signal processing means further comprising means for detecting differences between successive augmented frames of data corresponding to motion of the second image and for communicating augmented difference information to the computer to reposition display of the items through operating system controls, the items being repositioned on the display by amount corresponding to motion of the first and second images (figure 1 and 3a,b,c and column 5, lines 10-34). In regards to claim 9 Platzker shows a DSP for implementing select algorithms on difference frames or raw frames of image data (abstract and figure 3a and 6a). In regards to claim 10 Platzker (Continued on Supplemental Sheet.)

Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Sheet 10

Continuation of: Boxes I - VIII

V. 1. REASONED STATEMENTS:

The report as to Novelty was positive (YES) with respect to claims 6-8, 12-13, 15-29, 31-34, 36, 40-49, 52, 56-58, 63, 65-67, 69-71, 73-76.
The report as to Novelty was negative (NO) with respect to claims 1-5, 9-11, 14, 30, 35, 37-39, 50-51, 53-55, 59-62, 64, 68, 72, 77-81.
The report as to Inventive Step was positive (YES) with respect to claims 12, 22-23, 32-34, 36, 57, 65-67.
The report as to Inventive Step was negative (NO) with respect to claims 1-11, 13-21, 24-31, 35, 37-56, 58-64, 68-81.
The report as to Industrial Applicability was positive (YES) with respect to claims 1-81.
The report as to Industrial Applicability was negative (NO) with respect to claims NONE.

V. 2. REASONED STATEMENTS - CITATIONS AND EXPLANATIONS (Continued):

does not directly show his controller using specific memory type such as EPROM and RAM but examiner considers such common memory types to be inherent to the normal operation of a personal computer as shown in figure 1 of Platzker. In regards to claim 11 it is inherent that the DSP of Platzker is interfaced to the PCI bus in the computer shown in figure 1. In regards to claim 14 Platzker teaches frame memory to buffer one or more image frames (abstract). In regards to claim 30 Platzker show magnification in a projected image (figure 1). In regards to claim 35, 54, 55 Platzker show a second camera (figure 1). In regards to claim 39 Platzker shows a CPU figure 1. In regards to claims 37 and 38 Platzker does not directly state he has a DSP in his camera but he does state he has a CCD and a appropriate signal processing apparatus to the CCD therefore examiner contends a DSP is inherent to the camera. In regards to claims 50, 51 and 53 Platzker shows isolating one or more objects or symbols held by user (figure 3c). In regards to claim 59 Platzker does not directly claim his video camera has a transducer for converting optical signals to electrical signals but examiner contends this feature is inherent in an electronic video camera which uses a CCD (column 5, lines 45-57). In regards to claims 60-61 the computer comprises a digital signal processor separate from a CPU within the computer (column 5, lines 10-22). In regards to claim 62 Platzker teaches wherein the transducer, electronic means and signal processor are constructed and arranged into a single device in communication with the computer (figure 1 items 28a and 12b). In regards to claim 68 Platzker shows the transducer comprises one of a CCD or a CMOS integrated circuit with digital outputs (column 5, lines 45-57). In regards to claim 72 Platzker teaches wherein the signal processor comprises a video frame memory (column 5, lines 31-32). In regards to claim 78-81 Platzker teaches segmenting video images (figures 3b, 3c and 4, column 7, lines 22-29), image acquisition of portion of human hand and multiple features of a human (figures 1, 3b, 3c and 4) and a neural net means for learning user motion over time (figure 4, column 6, lines 40-49).

Claims 6-8, 12, 15-21, 24-29, 31, 40-49, 52, 56, 58, 63, 69-71, 73-76 lack an inventive step under PCT Article 33(3) as being obvious over Platzker et al. (5,528,263).

In regards to claims 6-8, 13, 15, Platzker does not directly show the difference electronics comprise multiple frame memory, a subtraction circuit, and state machine controller/memory addresser to control data flow, for storing and subsequently subtracting pixel-by-pixel difference data, but Platzker does state that he performs these functions and therefore it would have been obvious that he has difference electronics as broadly claimed which perform these functions (abstract and figures 6a and 6b, column 5, lines 10-34).

In regards to claims 16-18, 24-26, 73-76 Platzker suggest a signal processor with correlation means, best fit algorithm means and a peak detect algorithm, difference functionality and a two dimensional inverse FFT operation because these are common well known mathematical ways of detecting patterns and figure 3a item 46 suggest this (column 2, lines 35-53 and column 6, lines 34-49).

In regards to claims 19 and 20 Platzker does not directly show video cursor control being disabled but examiner contends that Platzker just illustrates a program and hardware connected to a PC and it would be inherent that one has the ability to turn program off and go back to convention mouse control and other keystrokes at the computer.

In regards to claims 21, 27-29, 31, 40-46, 49, 52 Platzker does not directly show controller responsive to a blink of a eye, movement of head of user but he does show controller responsive to hand movements and examiner contends that the Platzker apparatus has the capability to be also responsive to a blink of a eye and movement of the head and because the principle of operation of Platzker does not change ie detection of a body part movement such as hand. This modification would give the user more ways to implement more control options or functions and therefore it would have been obvious the modification of the Platzker apparatus to also detect movement of the head or eye.

In regards to claims 47 and 48 Platzker does not directly state that his camera has zoom and automatic focus but examiner contends that these features are well known in the prior art and since they are clearly desirable well known common

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/US99/00086

Supplemental Box

(To be used when the space in any of the preceding boxes is not sufficient)

Sheet 11

Continuation of: Boxes I - VIII

features on cameras which can make the camera more easy to use, which is motivation and therefore it would have been obvious to have these features on the Platzker camera. Note the disclosure does not teach how to implement a zoom or automatic focus and therefore it was reasonable to believe applicant was in agreement that such features in them self were well known at the time of his application and therefore it is only a question of was there known motivation at the time of invention to use well known camera features such as zoom and automatic focus in the Platzker camera.

In regards to claims 56 and 58 Platzker does not directly teach a re-calibration means as broadly claimed but examiner contends the positive benefits of calibration are well known in the prior art and that Platzker would want to keep his equipment in calibration to get accurate results which is the motivation to do re-calibration as broadly claimed.

In regards to claims 63 and 69-70 Platzker does not directly state that his cameras transducer generates a standard RS170 digital output for communication means to the computer interface. The examiner contends it would have been obvious to use the well known standard camera output such as RS170 because nonstandard outputs would be more costly and require more engineering modifications.

In regards to claim 71 Platzker does not directly state that the transducer generates digital resolutions of 4 bits or greater but examiner contends it would have been obvious that the CCD in Platzker would of had at least an 8 bit resolution which was the common video camera of the time which corresponds to the very common 256 gray scale.

Claims 12, 22-23, 32-34, 36, 57, 65-67 meet the criteria set out in PCT Article 33(2)-(4), because the prior art does not teach or fairly suggest in regards to claim 12 a MPEG compression electronics, in regards to claims 22-23 a video cursor control means responsive to sound generated by user, in regards to claim 32-34 a IR means for detecting heat associated with the user and for repositioning the items at a rate correlated to the heat, in regards to claim 36 a second camera that takes image data in a second electromagnetic spectrum, in regards to claim 65 the transducer is a visible CCD and IR camera, in regards to claims 66-67 CCD camera has 2x2 imaging pixels, in regards to claim 57 the re-calibration means comprises a microphone and the event comprises a sound generated by the user.

----- NEW CITATIONS -----

US 5,528,263 (PLATZKER ET AL.) 18 June 1996, abstract, figure 1, 3a, 5a, 6a, 6b, column 1, lines 23-68

HUMAN MOTION FOLLOWING COMPUTER MOUSE AND GAME CONTROLLER

Background

The primary human-computer interface employed today uses both a keyboard to enter textual information and a mouse to provide control over graphical interfaces presented to the user. These are the most widely used interfaces for computer applications which include word processing, presentation software (i.e. PowerPoint), computer aided design packages (i.e. AutoCAD), spreadsheet analysis (Excel) and others. These interfaces are also widely used for computer gaming entertainment, though often augmented by or replaced by a joystick.

In the daily use of business applications, the need to control the cursor position on the screen requires that the user remove his/her hand from the keyboard in order to use the standard mouse. The use of the mouse introduces several issues. In a desk environment, the use of the mouse requires maintenance of a free area on the desk. The mouse cord must also remain free from obstruction in movement. Additionally, the use of the mouse is a major contributing factor for carpal tunnel syndrome. It would be advantageous therefore to find an alternative to the mechanical mouse.

In computer gaming, the complexity of the controls required to operate the game requires a combination of either mouse and keyboard or joystick and keyboard. Gaming application require the user to control many axis of motion. These can include forward motion, reverse motion, left turn, right turn, left strafe (slide), right strafe, upward motion, downward motion. Additionally, many games allow the user to look in directions different from that in which the vehicle is moving including up, down, left and right. These many axes of motion therefore drive the complexity and enjoyment of the game.

It is also desirable to offer alternative approaches to human-computer interfaces to those incapable of using standard devices due to disability.

The object of this invention is to provide the means to either replace or augment the existing human computer interfaces by allowing the operator to control the cursor position by motion of the user observed by a video camera. This motion may be imparted by the up-down left-right motion of the user's head, hands or other motion presented to the video camera. In a close up view of the users facial features, the up down, left-right motion may be imparted through rotation as the facial features observed by the camera will appear as a translation. While systems are available to achieve this capability, they remain prohibitively expensive to the general public. These costs are driven by the techniques and algorithms by which these systems detect the user head motion. Additionally, all techniques in existence today require the user to augment the system by wearing a detectable target or an apparatus which emits or detects a signal, all of which are cumbersome to the user. The unique element in this invention which brings the system cost into the range of the general public is a unique and efficient algorithm for

detecting user motion without the aid of augmentation by artificial devices placed on the operator.

One object of this invention is to provide a means of human control of a graphical computer interface through the physical motion of the user in order to control the activity of a cursor in the manner usually accomplished with a computer mouse.

A further object of this invention is to provide additional degrees of freedom in the human computer interface in support of computer games and entertainment software.

A further object of this invention is the dual use of the electronics and camera used for human computer interface in support of teleconferencing and video frame capture.

Summary of Invention

The invention provides the means for a human to interface to a computer via physical motion of the user, i.e. the users head or hands. It also provides a human factors approach to cursor movement in which the users rate of motion determines the relative motion of the cursor i.e. the faster the users head travels over set distance the further the corresponding cursor movement.

In one aspect, the invention includes an electro-optic sensor consisting of an array of imaging elements. This sensor may be either a visible light camera utilizing ambient lighting conditions or a camera sensitive in another band such as the near IR, in which the illumination source is from an IR lamp which is beyond human sensory perception. This sensor is mounted facing the user in such a way that the user's face is captured by the sensor.

The images captured by the sensor are processed by a digital signal processor located either in the users computer. The detected motion of the user is communicated to the user's operating system via the pci bus interface. These commands are interpreted by a low overhead program operating on the users main processor that either updates the cursor position on the screen or provides motion information to the users computer game.

Alternatively, the digital signal processor may be mounted in the camera housing such that the camera/signal processing subsystem produces signals which emulate the mouse via the mouse input connector.

The pixel format of the camera drives the accuracy of the system. It should be noted that the greater the density of pixels on the image of the users face, the higher the resolution of cursor motion which can be attained. Camera formats of 240 vertical by 320 horizontal should provide satisfactory performance. The number of pixels that may be utilized is strictly determined by system cost factors. The greater the number of pixels, the more powerful the DSP must be in order to process the image sequences in real time. Current technology limits the processing density to a 64x64 window for a consumer.

While this is satisfactory for the general user, a higher fidelity system using a greater number of pixels is possible, at a proportionally higher cost.

The data transfer rate from the camera is 30 frames/second at 240x320 pixels per frame. Assuming eight bits per pixel, the digital data transfer rate is therefore 18.432 megabits/second. This is a high transfer rate for a consumer product using current technology. While the data transfer can be either analog or digital, the preferred method of image data transfer for this invention is therefore via a standard RS170 analog video interface.

Brief Description of the Drawings

Figure 1 illustrates the human computer interface according to the invention

Figure 2 illustrates the functions required on the printed circuit card

Figure 3 illustrates the algorithm that forms the basis of the invention

Detailed Description of the Drawings

Figure 1 illustrates the major components of the human computer interface. The user 5 sits facing the computer display 10. A camera 15 is mounted to the monitor facing the computer user 5. The camera 15 is mounted in such a way that the user's face is imaged. The camera 15 is interfaced to a printed circuit card 20 mounted in the users computer chassis 25. The camera 15 interfaces to the printed circuit card 20 via a camera interface cable 30.

Figure 2 illustrates the functions required on the printed circuit board 20 mounted in the users computer 25. A camera interface circuit 50 receives the video data. In the preferred implementation, this data is RS170 format. This function decodes the analog video data to determine the video timing signals embedded in the analog data. These timing signals are used to control an analog to digital converter that converts the analog pixel data in to digital camera images. The analog data is digitized into 6 bits in the preferred implementation although any number of bits greater than six may be acceptable and/or required for other optional features on the printed circuit card.

The frame difference electronics 55 receive the digital data from the camera interface circuit 50. The frame difference electronics include a single frame memory, a subtraction circuit and a state machine controller/memory addresser to control the data flow. The frame memory holds the previous digitized frame. As each digitized pixel is received by the frame difference electronics, the corresponding pixel from the previous frame is read from the frame memory and subtracted from the current. The resulting difference is output to the N frame video memory. The new frame pixel data is then stored into the frame difference electronics frame memory.

The N frame video memory electronics 60 receives either the differenced frames output by the frame difference electronics 55 or the raw digitized frames from the camera interface 50. The choice of frame source is made by the software resident on the digital signal processor 65. This frame video memory is sized to hold greater than one full frame of video up to a number of frames, N. The number of frames is to be driven by the final hardware and software design.

The digital signal processor 50 implements the algorithm unique to this invention. This algorithm determines the rate of head motion of the user in two dimensions. The digital signal processor also detects the eye blink of the user in order to emulate the click and double click action of a standard mouse button. In support of these functions, the digital signal processor commands the N frame video memory 60 to supply either the differenced frames or the raw digitized frames. The digital signal processor requires a supporting program memory 70 made up of electrically reprogrammable memory (EPROM) and data memory consisting of standard volatile random access memory (RAM). The digital signal processor is also provided with an interface to the PCI bus interface electronics 80 through which the cursor and button emulation is passed to the users main processor. The PCI interface also provides the means to pass raw digitized video to the main processor as a user optional feature. This same interface provides the means, via the digital signal processor interface, to reprogram the program memory 70, allowing for software upgrades to provide additional features and performance.

The PCI interface electronics provide an industry standard bus interface supporting the aforementioned communication path between the printed circuit card 20 and the users main processor 25.

The optional MPEG compression electronics 85 provide the printed circuit card and camera to provide compressed video to the users main processor. This compressed video supports the use of the human computer interface electronics and camera in teleconferencing applications. This dual use as either human computer interface and teleconferencing is a unique combination offering the user an economical solution to two distinct applications.

Figure 3 describes the head motion algorithm upon which this invention is based. Note that not all of the functions shown in Figure 3 are implemented in software in the DSP. As with other algorithms typical of the detection of motion, this algorithm relies on the correlation of images from one frame to the next. The unique aspect of this invention lies in the use of frame differenced images in the correlation process. The frame differencing operation removes all parts of the camera images that are unchanged from the previous frame. All room background behind the user is therefore removed from the image. This greatly simplifies the detection of feature motion. The users face image consists of regions of uniform illumination such that even with the users facial motion, these uniform regions (i.e. cheeks, forehead, chin) will also be removed. Note the users face also consists of dynamic features such as nose, eyes, eyebrows and mouth which have enough

spatial detail that they will be evident in the differenced image. As the user moves his face with respect to the room lighting, the shape and distribution of these features will change, but the frame rate of the camera ensures that these features look similar from one frame to the next. The correlation process therefore operates to determine how these differenced features are moving from one frame to the next in order to determine user head motion.

The algorithm receives video images of the user, imaged over time. Each image received is provided to both a frame memory 100 and a differencer 105. Though the preferred implementation is to buffer a single frame in this memory, the memory may consist of many frames, buffered such that the first frame input is the first frame output (FIFO). The delayed frame is read from the frame memory 100 and subtracted from the current frame using the differencer 105. The frame output is provided to both a correlation process 115 and a difference frame memory 110.

Like the frame memory, while the preferred implementation requires only a single difference frame, the difference frame memory 110 can hold many difference frames in sequence in a FIFO arrangement. The delayed difference frame is read from the difference frame memory and provided to the correlation function 115. The correlation process 115 determines the best combination of row and column shifts in order to minimize the difference between the current difference frame and the delayed difference frame. The number of rows and columns required to align these difference images provides information on the users motion. The best-fit function 120 determines the row and column shift giving the optimum alignment. In the case of a classical correlation process, the best-fit function 120 would consist of a peak detect.

The best-fit function 120, provides the relative motion in rows and columns of the observed users features. The cursor update compute function 125 translates this measured motion into the position change required of the cursor. This is likely to be a non-linear process such that with greater head motion, the cursor moves a non-proportionally greater distance. For example a 1-pixel user motion will cause the cursor to move one screen pixel while a 10-pixel user motion may cause a 100-pixel screen cursor motion.

The video cursor control 135 provides a user interface to enable and disable the operation of the video human motion cursor control. This control is implemented either through a combination of keystrokes on the users keyboard, by sensing the eye-blink of the user or through voice commands. The functionality of the video cursor control 135 is to provide the user with the equivalent of a mouse pick-up, put-down action. As the user move the cursor from left to right across the screen, the user would de-activate the motion based cursor control in order to allow the user to move his head back to the left. Once the user has recentered his head, the user would once again activate the cursor control and continue to move the cursor about the screen. The activation/deactivation of the mouse input is represented by the switch 140, such that the open position of the switch disables

the human motion control of the cursor and supplies a zero change input to the summation operation 130.

When video cursor control 135 enable the human motion control of the cursor, the result of the cursor update compute function 125 is added to the known current cursor position by the summer 130. This summation has an x component and a y component. The result of the summation 130 is used to update the cursor position on the users screen via the users operating system. The current cursor position is provided by the computer operating system. The cursor position may be controlled by both the user visible motion as well as the motion imparted by other input device such as a standard computer mouse.

Figure 4 provides a detailed description of the preferred implementation of the algorithm described in functions 100-120 in Figure 3.

Video data is received by the processing electronics in both a single frame memory 200 and a differencer 205. The output of the frame memory 200 is also provided to the differencer 205 such that the previous frame is subtracted from the current frame. This differenced frame is then processed by a two dimensional FFT 210.

The complex result of the FFT 210 is provided to both a complex multiplier 225 and a complex memory 215. The complex memory is the size of the processed image, each location containing both a real and imaginary component of a complex number. With each new FFT operation 210, the previous FFT result, contained in the complex memory 215, is provided to the conjugate operation 220. The complex conjugate of each element is computed and provided to the complex multiplier 225. In this manner, the FFT of the previous frame difference is conjugated and multiplied against the FFT of the current difference image.

The two dimensional array of complex products output by the complex multiplier 225 is provided to a two dimensional inverse FFT operation 230. This operation creates an image of the correlation function between the latest pair of difference images. The correlation image is processed by a peak detection function in order to determine the shift required aligning the two difference images. The x-y magnitude of this shift is representative of the users motion. This x-y magnitude is provided to the software that will be used to update the cursor position as described in Figure 3.

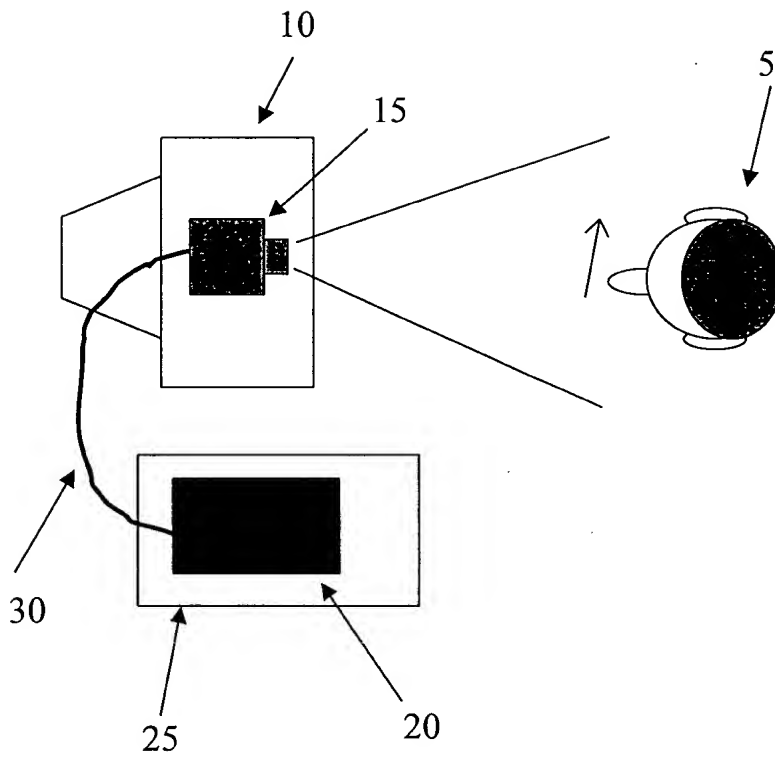


Figure 1

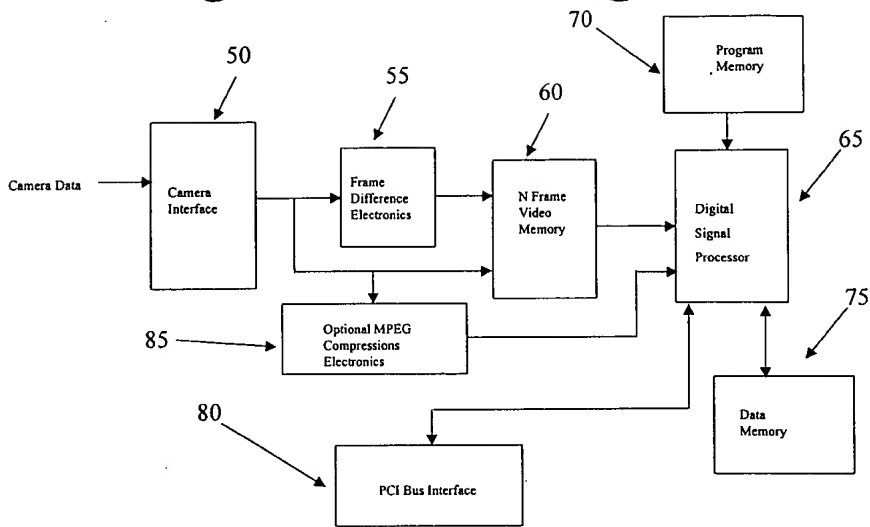


Figure 2

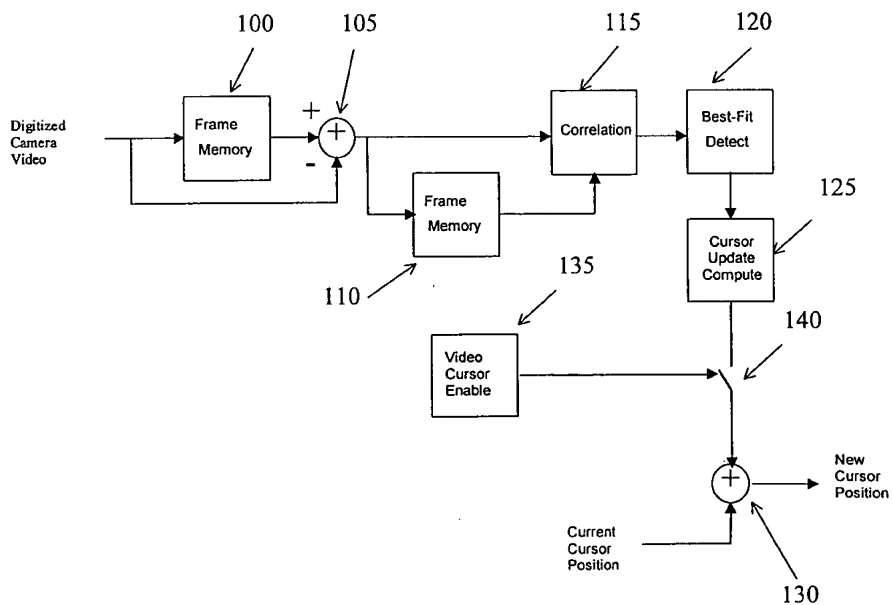


Figure 3

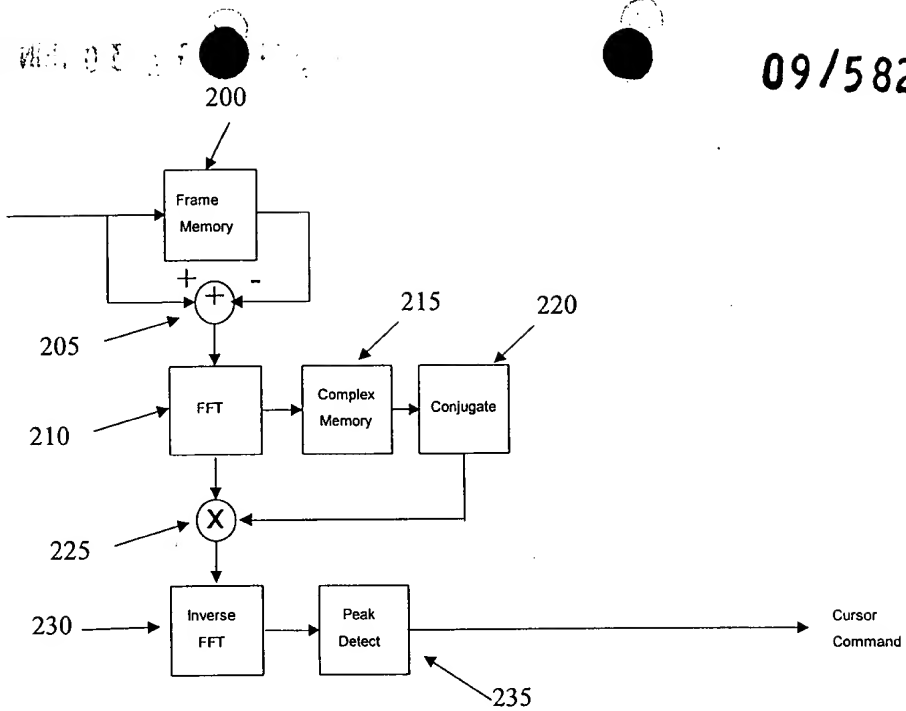


Figure 4

1/20

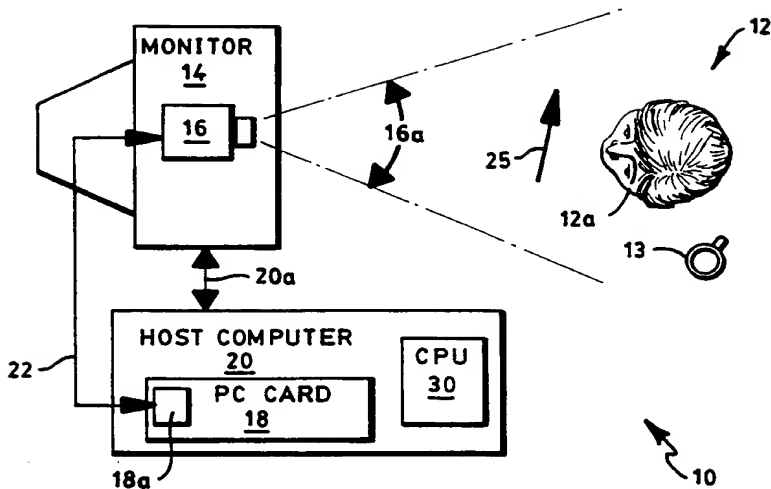


FIG. 1

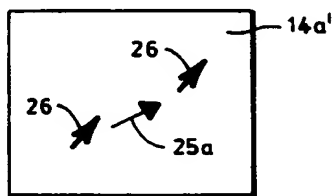


FIG. 1A

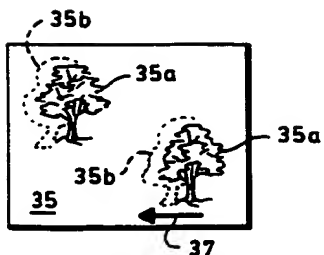


FIG. 1B



FIG. 1C

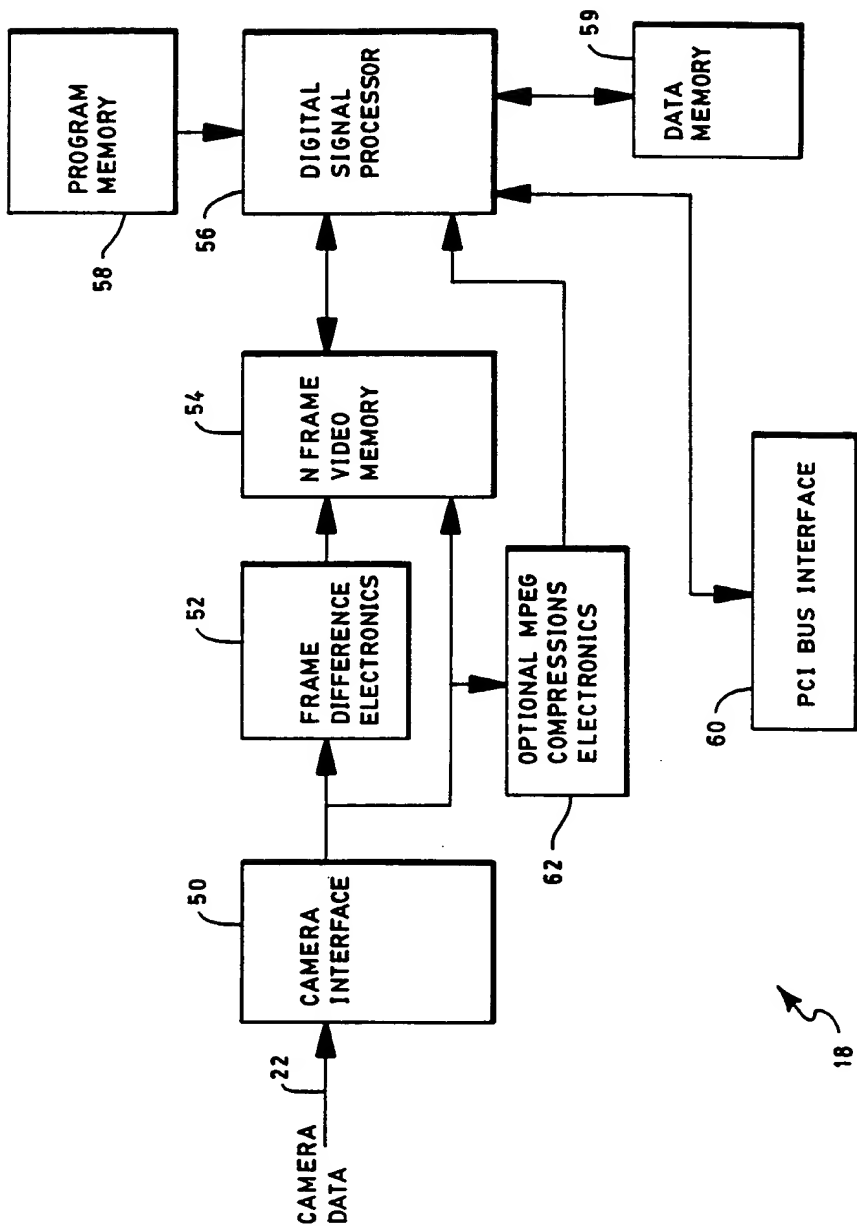


FIG. 2

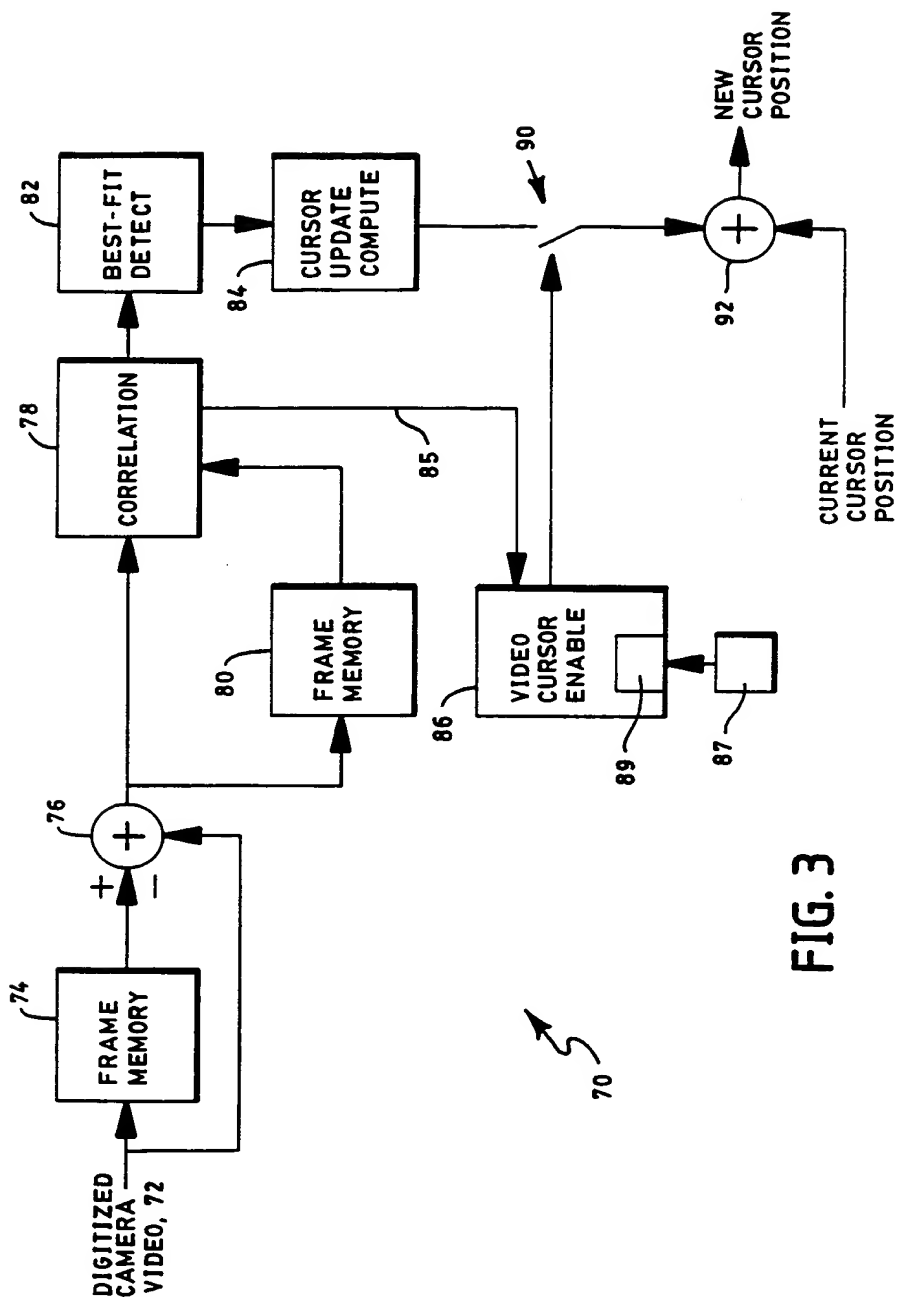


FIG. 3

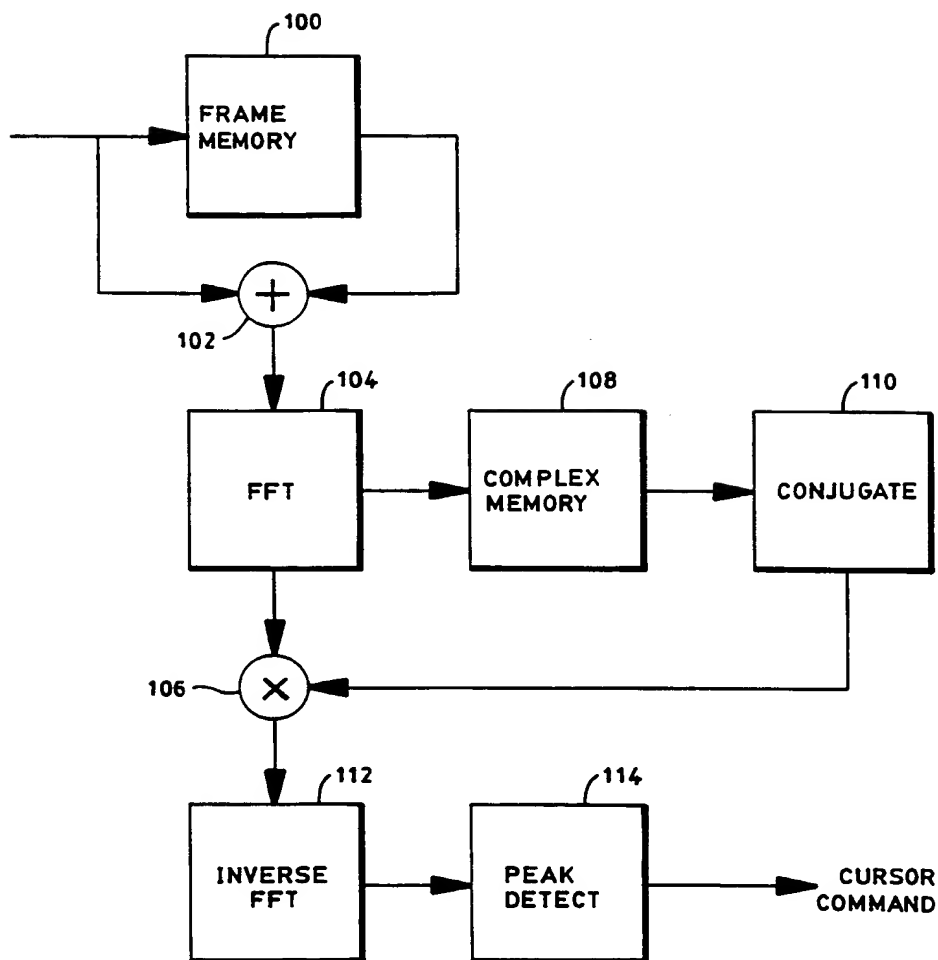


FIG. 4

5/20

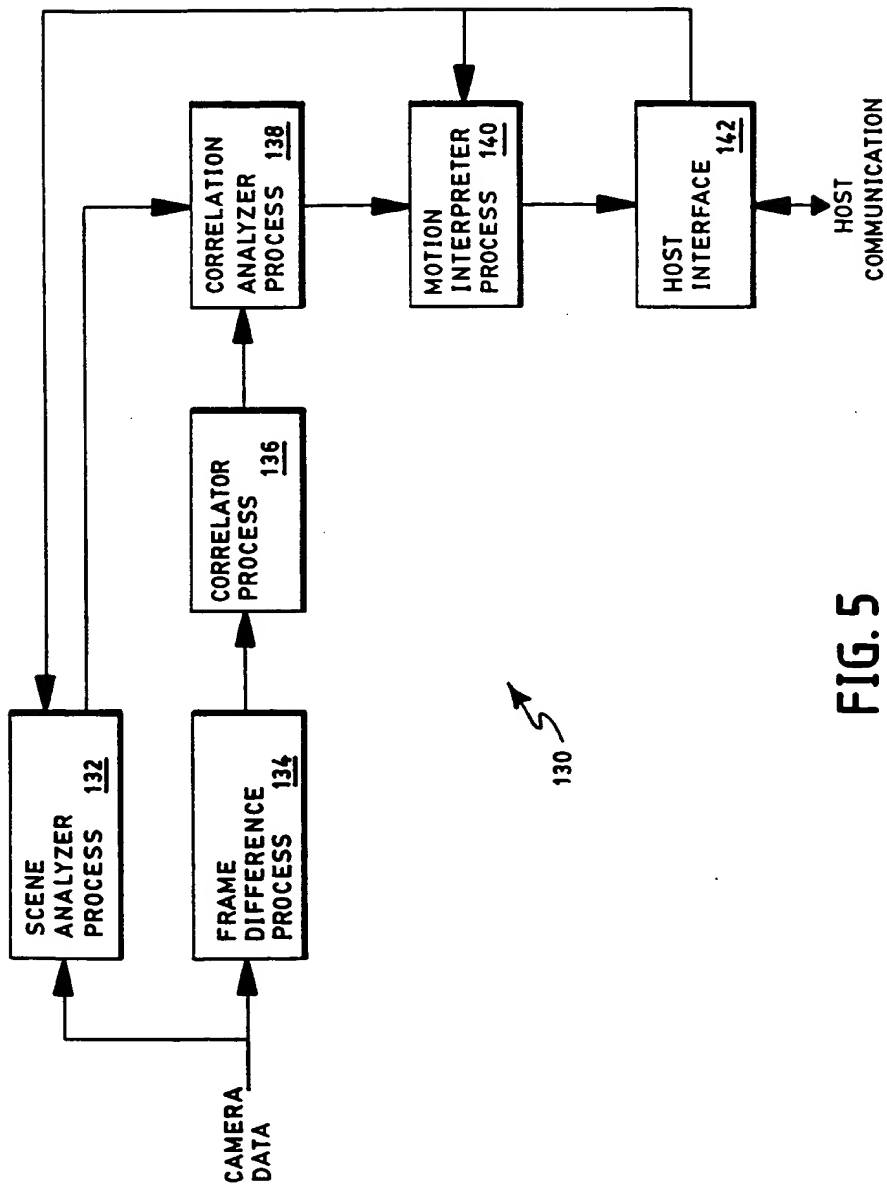


FIG. 5

6/20

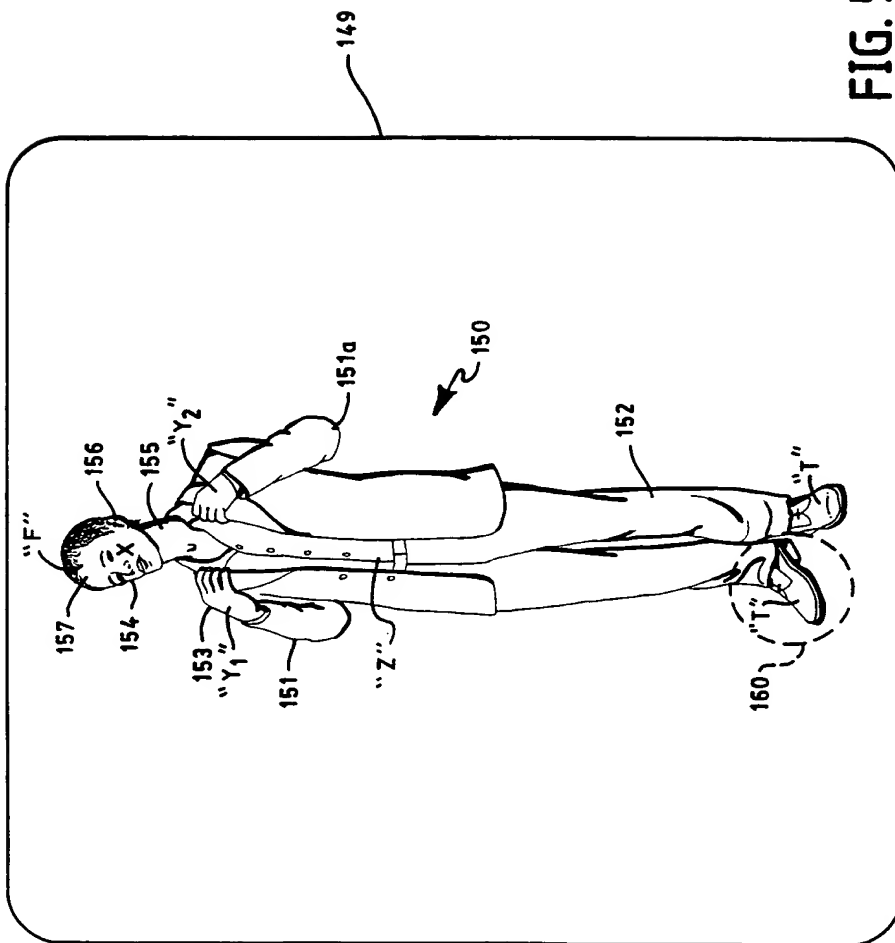


FIG. 5A

7/20

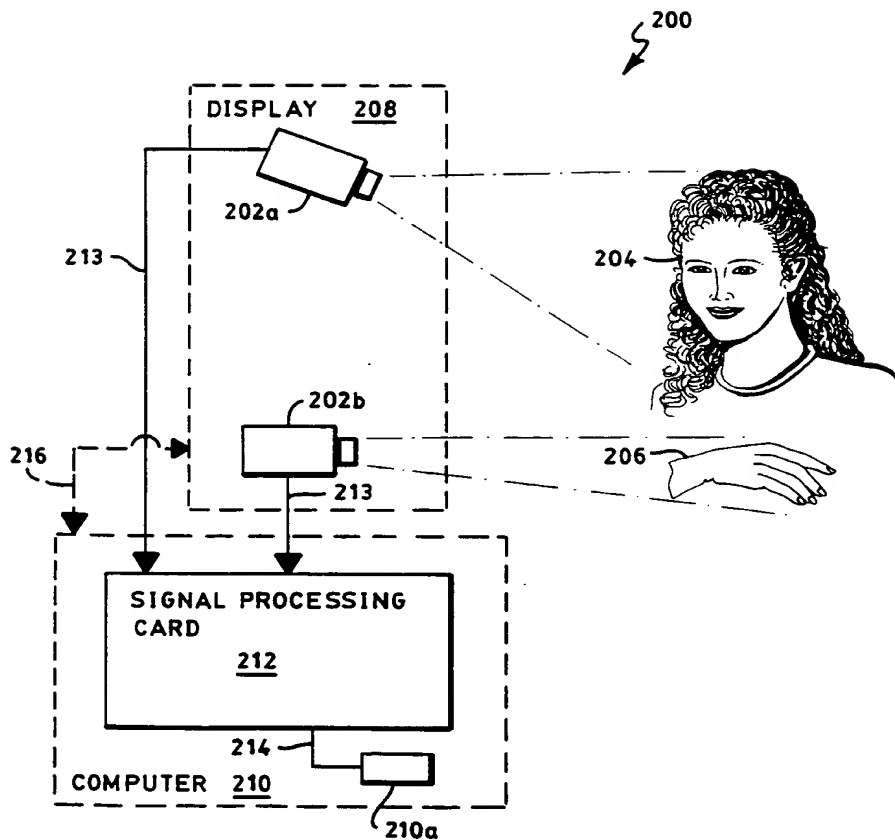


FIG. 6

8/20

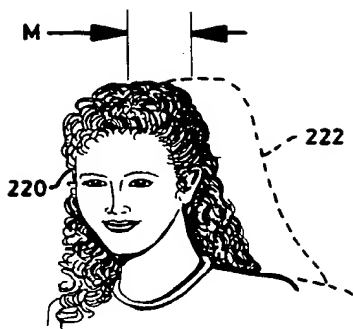


FIG. 7

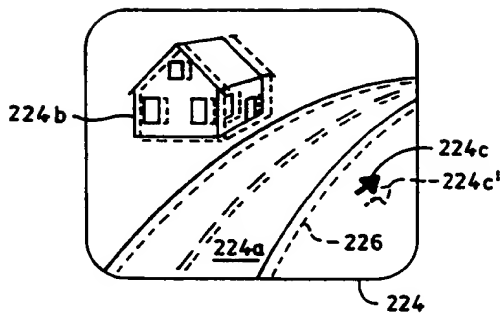


FIG. 7A

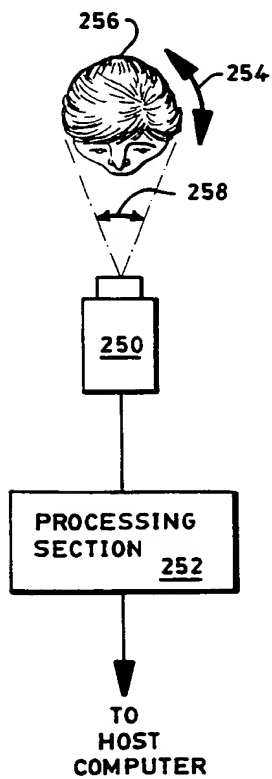


FIG. 8

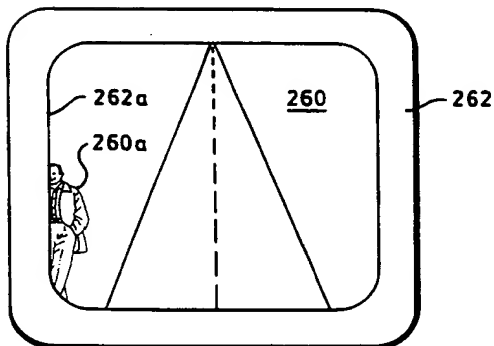


FIG. 8A

TO HOST
COMPUTER

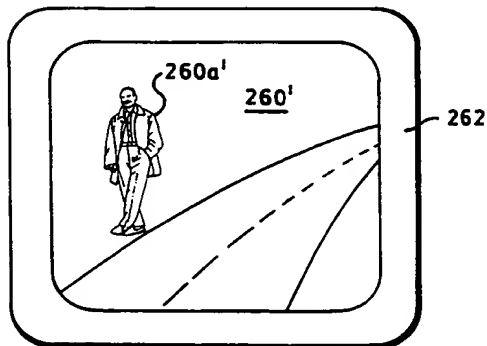


FIG. 8B

TO HOST
COMPUTER

9/20

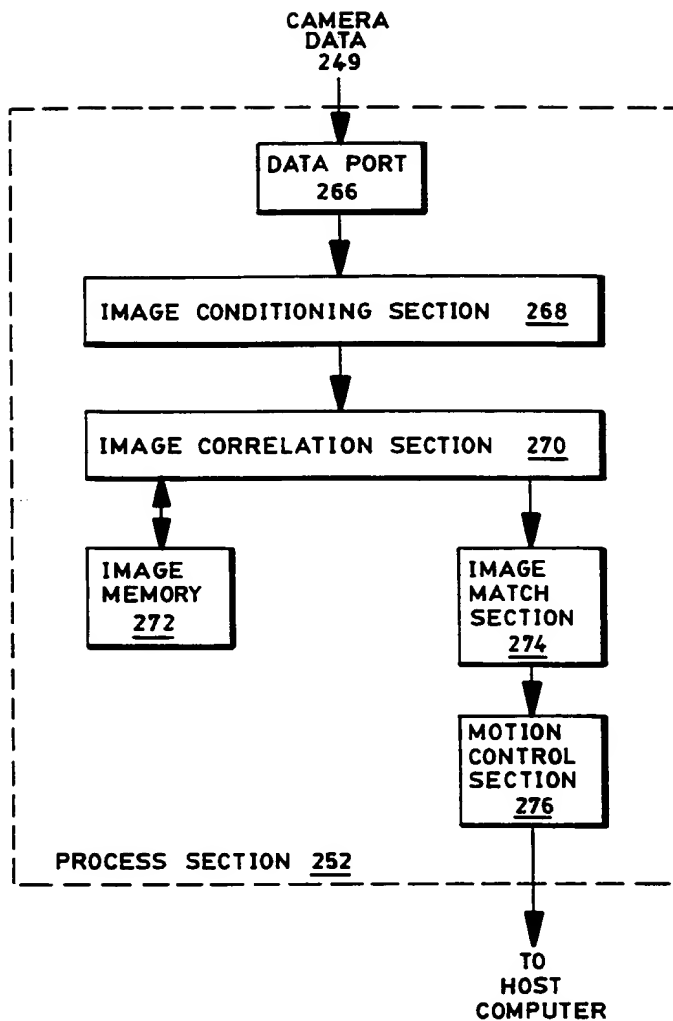


FIG. 8C

10/20

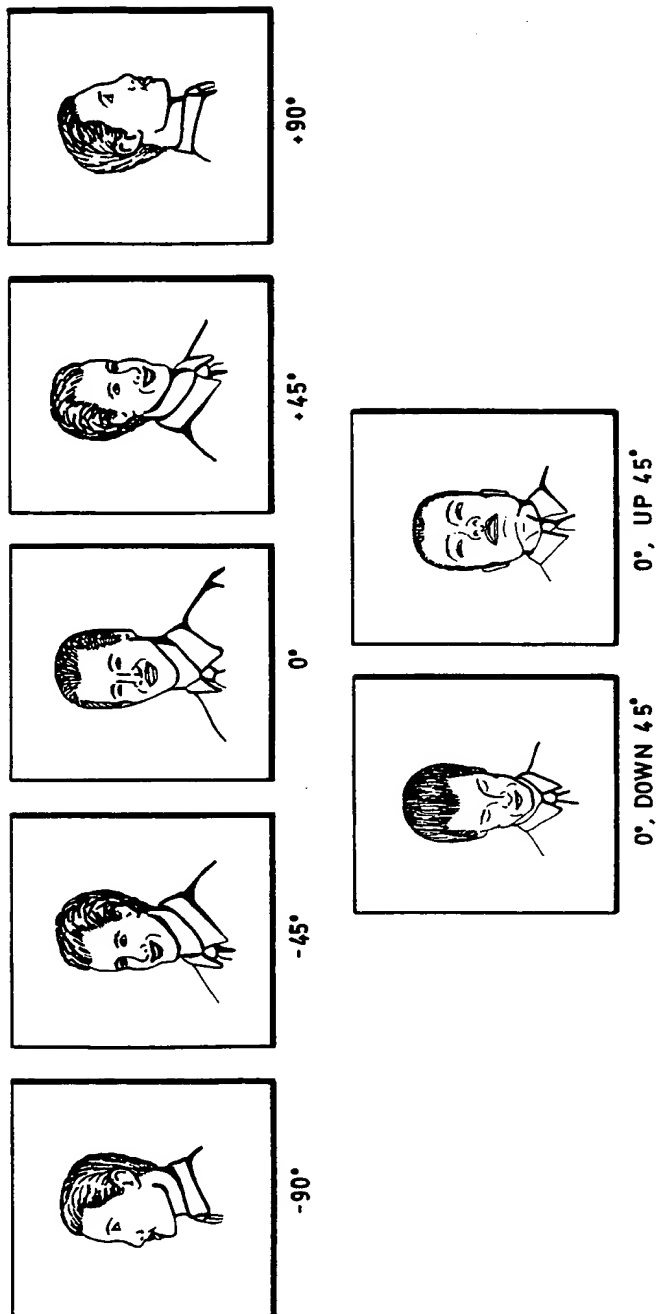


FIG. 8D

11/20

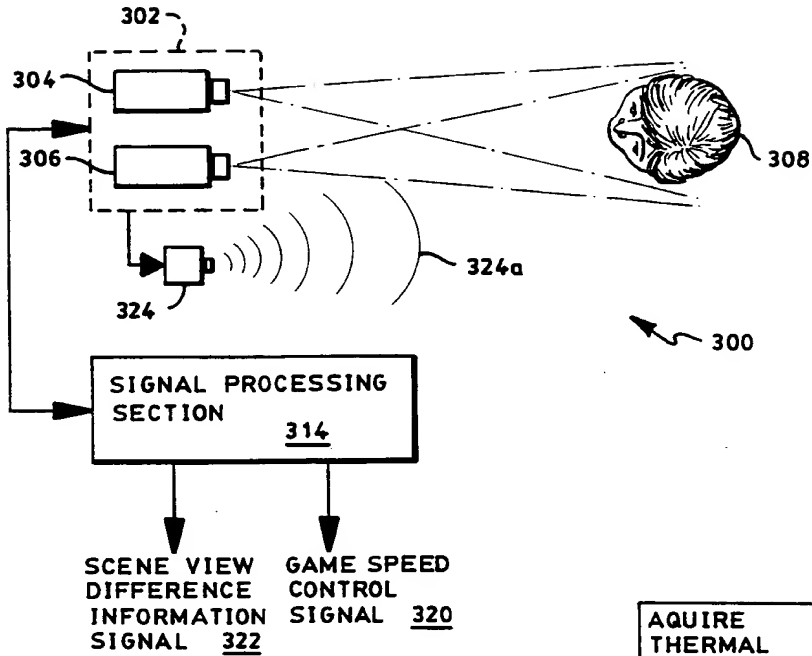


FIG. 9

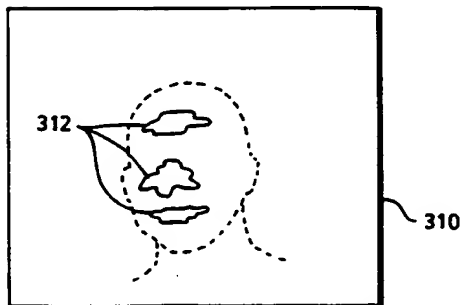


FIG. 9A

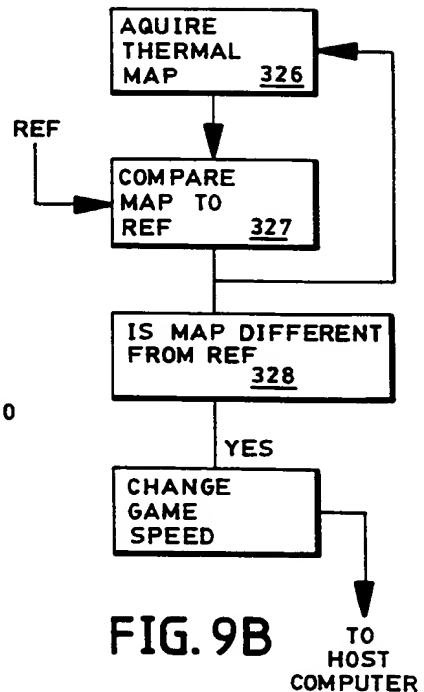


FIG. 9B

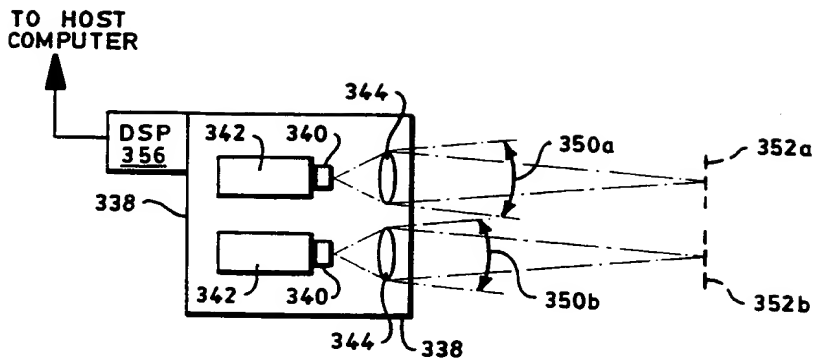


FIG. 10

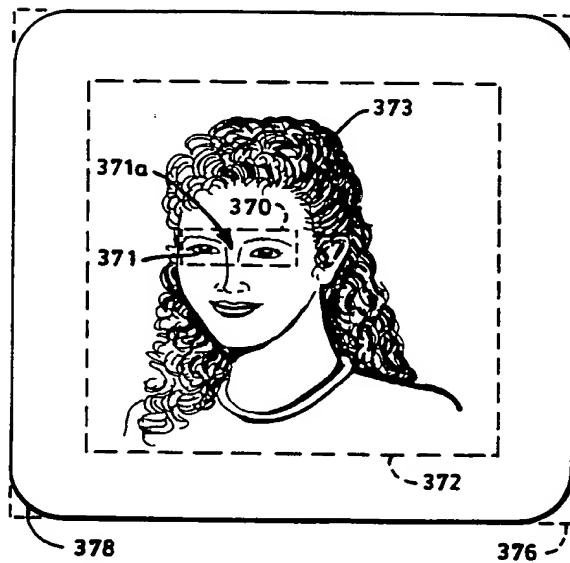


FIG. 11



FIG. 12

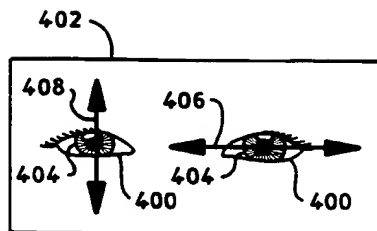


FIG. 12A

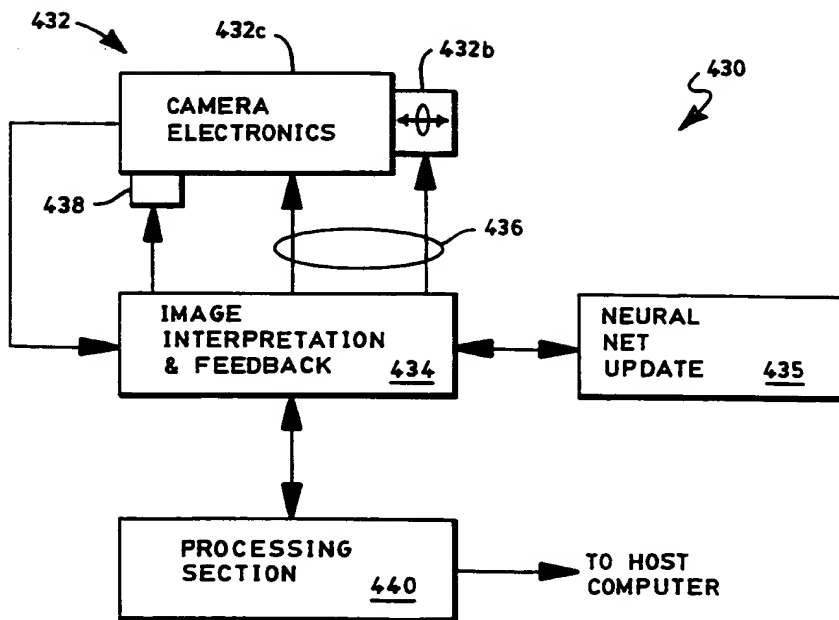


FIG. 13

14 / 20

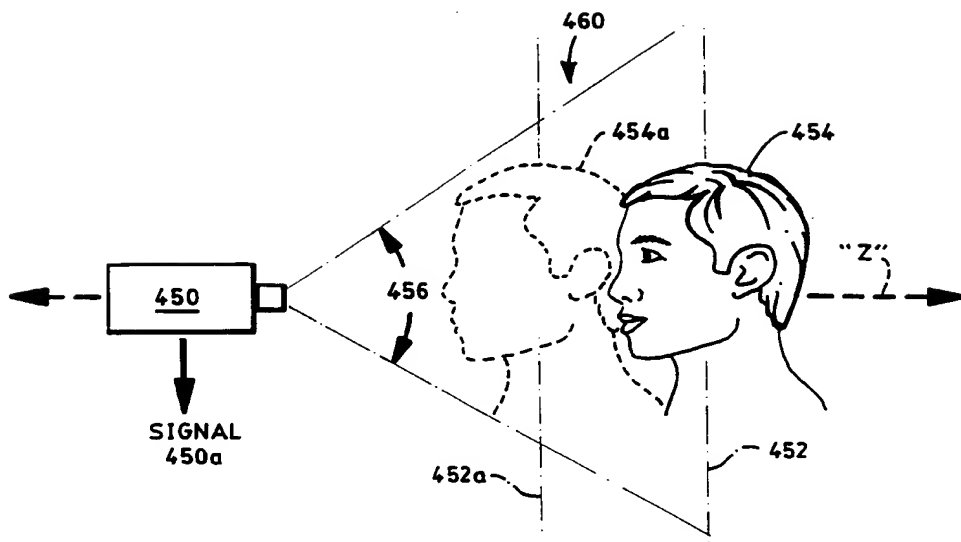


FIG. 14

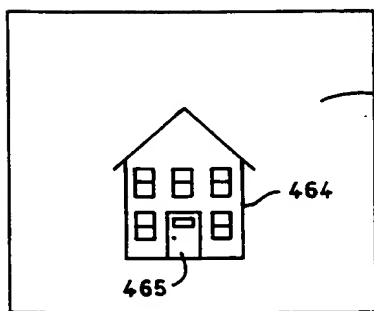


FIG. 14A

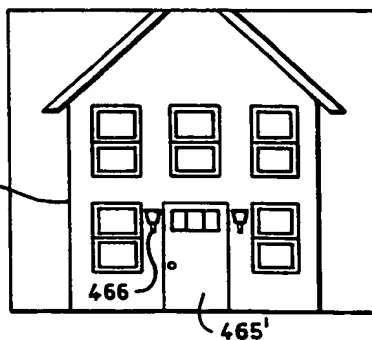


FIG. 14B

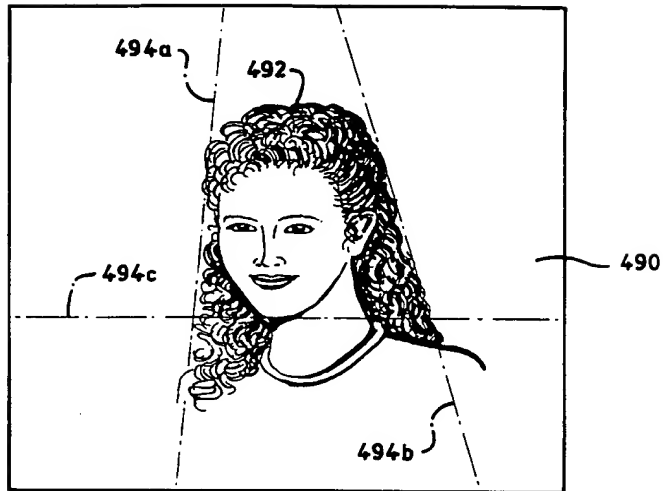


FIG. 15

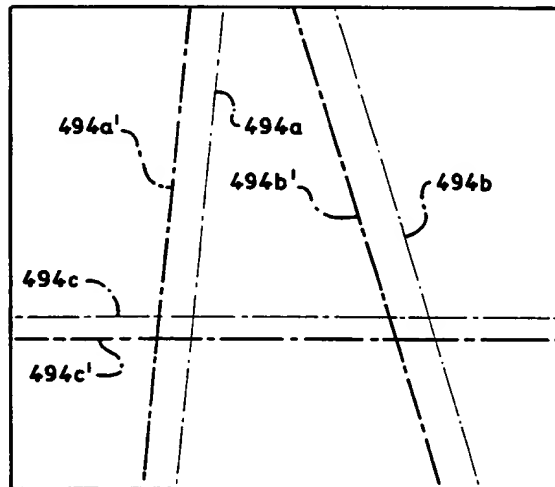


FIG. 15A

16/20

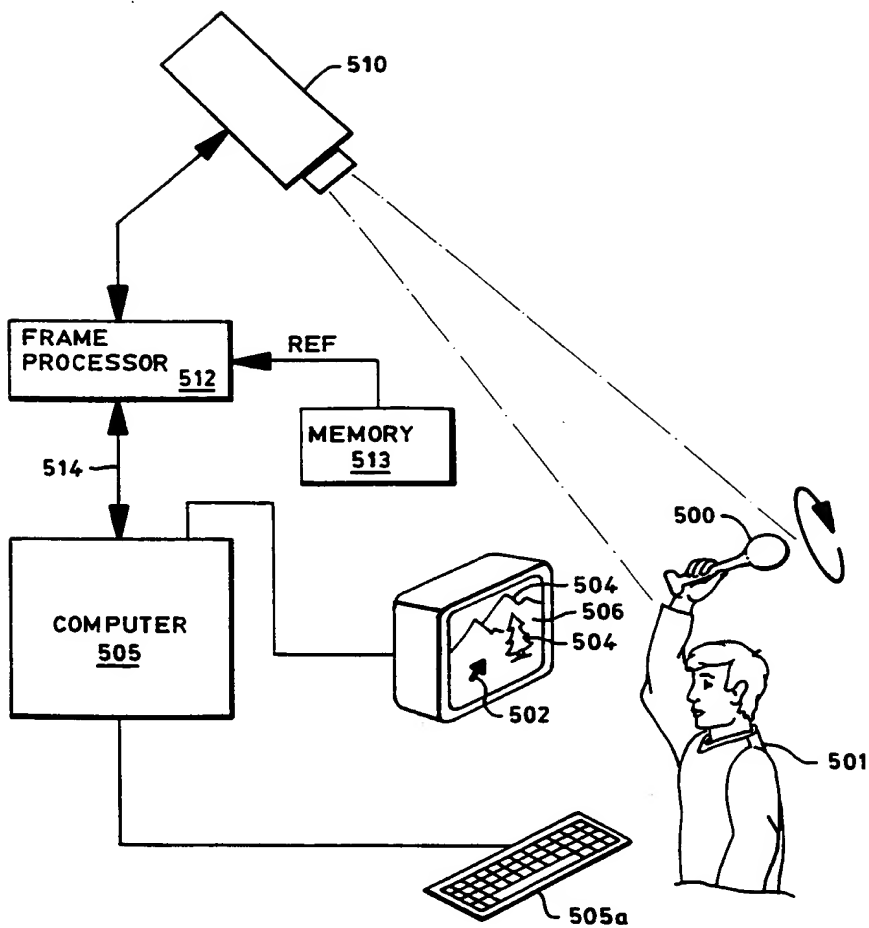
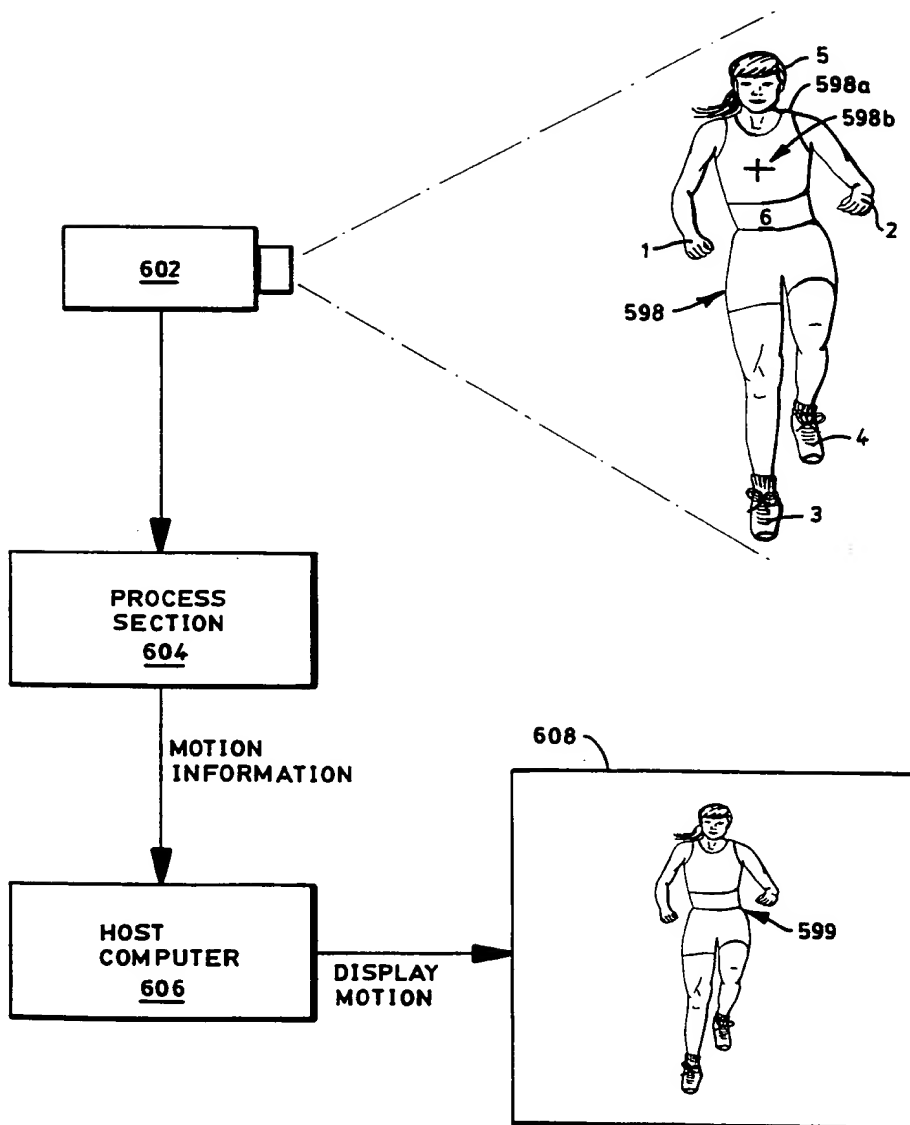


FIG. 16

17/20

**FIG. 17**

18/20

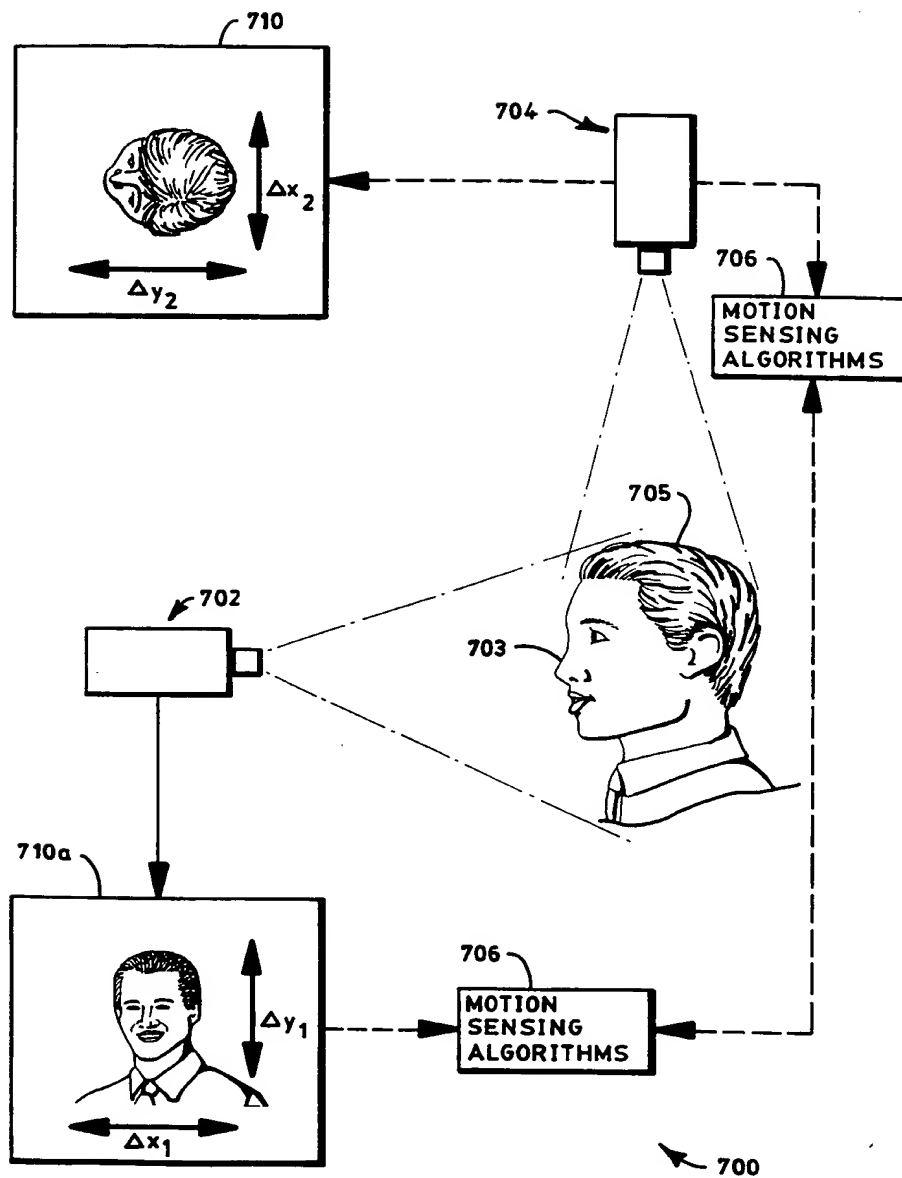


FIG. 18

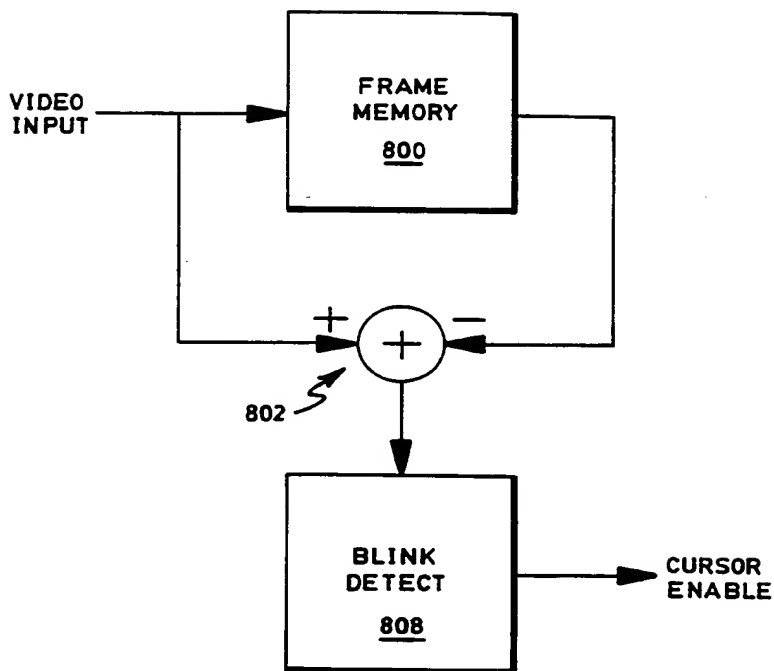


FIG. 19

20/20

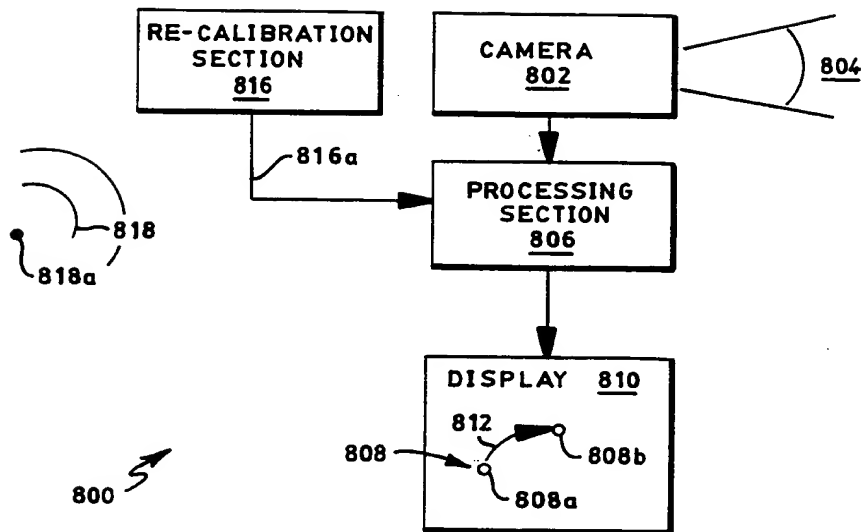


FIG. 20

09/582806

526 Rec'd PCT/PTO 30 JUN 2000

PATENT COOPERATION TREATY
RECEIVING OFFICE/UNITED STATES
UNITED STATES PATENT AND TRADEMARK OFFICE

International Application No.:)	
PCT/US99/00086)	
International Filing Date:)	
04 January 1999 (04.01.99))	
For: HUMAN MOTION FOLLOWING)	
COMPUTER MOUSE AND GAME)	
CONTROLLER)	Agent's Ref. No: 4521/001PC
Applicants:)	
THE VIDEO MOUSE GROUP et al.)	

RESPONSE TO THE INVITATION TO CORRECT DEFECTS
IN THE INTERNATIONAL APPLICATION AND TO COMMUNICATION
IN CASES FOR WHICH NO OTHER FORM IS APPLICABLE

Dear Sir:

This is in response to the Invitation to Correct Defects in the International Application and to the Communication in Cases for which No Other Form is Applicable, mailed 09 February 1999.

Applicants submit herewith two duly executed Powers of Attorney to complete the set of three (the Power executed by the first-named applicant was filed together with the application).

Applicants submit herewith substitute sheet 68 of the claims (1 sheet), on which claims 79-81 have been renumbered to correct the duplicate numbering of claim 79 in the application as filed, as pointed out by the RO/US in the Communication. Please substitute the enclosed sheet 68 for the corresponding sheet in the application as filed.

Although not requested by the RO/US, applicants also submit herewith substitute sheets 1/20-20/20 of the drawings (20 sheets), on which the overall quality of the drawings has been improved, especially on sheet 20/20 (Fig. 20), without adding any new subject matter. Please

- 2 -

substitute the enclosed sheets 1/20-20/20 of the drawings for the corresponding sheets in the application as filed.

Respectfully submitted,
DUFT, GRAZIANO & FOREST, P.C.

Date: 2/25/99

By: C. A. Vock
Curtis A. Vock (Reg. No. 38,356)
1790 - 30th Street, Suite 140
Boulder, CO 80301, USA
Tel: 303-449-9497
Fax: 303-449-0814

Enclosures

PCT

POWER OF ATTORNEY

(for an international application filed under the Patent Cooperation Treaty)
(PCT Rule 90.4)

The undersigned applicants (Names should be indicated as they appear in the request):

FREY, Robert D.
19 Still River Road
Bolton, Massachusetts 01740
United States of America

GREALISH, Kevin
18 Whitney Avenue
Westwood, Massachusetts 02090-2948
United States of America

MARSHALL, Charles M.
667 N. Bradford Street
North Andover, Massachusetts 01845
United States of America

VOCK, Curtis A.
1279 Elder Avenue
Boulder, Colorado 80304
United States of America

hereby appoint the following persons as ☒ agent ☐ common representative

Vock, Curtis A.; Duft, Donald M.; Graziano, James M.; Forest, Carl A.; Cleveland, Dan Jr.;
Setter, Michael J.; Wilbar, William P.; Swenson, Thomas; Kirk D. Williams
DUFT, GRAZIANO & FOREST, P.C.
1790-30th Street, Suite 140
Boulder, Colorado 80301-1018
United States of America

to represent the undersigned before ☒ all the competent International Authorities
☐ the International Searching Authority only
☐ the International Preliminary Examining Authority only

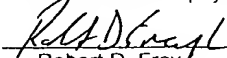
in connection with the international application identified below:

Title of the Invention: HUMAN MOTION FOLLOWING COMPUTER MOUSE AND GAME
CONTROLLER

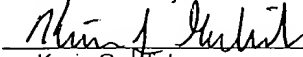
Applicant's or agent's file reference: 4521/001PC

International application number (if already available):

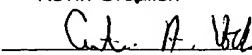
filed in the United States Patent and Trademark Office as Receiving Office on 04 January 1999, and
to make or receive payments on behalf of the undersigned.


Robert D. Frey

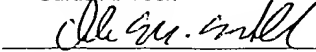
Date: 7 February, 1999


Kevin Greulich

Date: 17 February, 1999


Curtis A. Vock

Date: Jan. 4, 1999


Charles M. Marshall

Date: 17 February, 1999

PCT

POWER OF ATTORNEY

(for an international application filed under the Patent Cooperation Treaty)
(PCT Rule 90.4)

The undersigned applicants (*Names should be indicated as they appear in the request*):

FREY, Robert D.
19 Still River Road
Bolton, Massachusetts 01740
United States of America

GREALISH, Kevin
18 Whitney Avenue
Westwood, Massachusetts 02090-2948
United States of America

MARSHALL, Charles M.
667 N. Bradford Street
North Andover, Massachusetts 01845
United States of America

VOCK, Curtis A.
1279 Elder Avenue
Boulder, Colorado 80304
United States of America

hereby appoint the following persons as ☒ agent ☐ common representative

Vock, Curtis A.; Duft, Donald M.; Graziano, James M.; Forest, Carl A.; Cleveland, Dan Jr.;
Setter, Michael J.; Wilbar, William P.; Swenson, Thomas; Kirk D. Williams
DUFT, GRAZIANO & FOREST, P.C.
1790-30th Street, Suite 140
Boulder, Colorado 80301-1018
United States of America

to represent the undersigned before ☒ all the competent International Authorities
☐ the International Searching Authority only
☐ the International Preliminary Examining Authority only


in connection with the international application identified below:

Title of the Invention: HUMAN MOTION FOLLOWING COMPUTER MOUSE AND GAME CONTROLLER

Applicant's or agent's file reference: 4521/001PC

International application number (if already available):

filed in the United States Patent and Trademark Office as Receiving Office on 04 January 1999, and to make or receive payments on behalf of the undersigned.

Robert D. Frey	Date:	_____ , 1999
Kevin Grealish	Date:	_____ , 1999
 Curtis A. Vock	Date:	<u>Jan. 4</u> , 1999
Charles M. Marshall	Date:	_____ , 1999

PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

T. CURTIS A. VOCK
DUFT, GRAZIANO & FOREST, P.C.
1790-30TH STREET
SUITE 140
BOULDER CO 80301-1018

PCT

DUFT, GRAZIANO &
FOREST, P.C.

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT OR THE DECLARATION

(PCT Rule 44.1)

Date of Mailing
(day/month/year)

07 JUL 1999

Applicant's or agent's file reference

4521001PC

FOR FURTHER ACTION See paragraphs 1 and 4 below

International application No.

PCT/US99/00086

International filing date
(day/month/year)

04 JANUARY 1999

Applicant

THE VIDEO MOUSE GROUP

*Art. 19? 9/7/99
Comments on Abstract? 8/7/99*

Docketed for

1. ☒ The applicant is hereby notified that the international search report has been established and is transmitted herewith.

Filing of amendments and statement under Article 19:

The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

When? The time limit for filing such amendments is normally 2 months from the date of transmittal of the international search report; however, for more details, see the notes on the accompanying sheet.

Where? Directly to the International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland
Facsimile No.: (41-22) 740.14.35

For more detailed instructions, see the notes on the accompanying sheet.

2. ☐ The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. ☐ With regard to the protest against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

☐ the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

☐ no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Further action(s):** The applicant is reminded of the following:

Shortly after 18 months from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in rules 90 *bis* 1 and 90 *bis* 3, respectively, before the completion of the technical preparations for international publication.

Within 19 months from the priority date, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later).

Within 20 months from the priority date, the applicant must perform the prescribed acts for entry into the national phase before all designated Offices which have not been elected in the demand or in a later election within 19 months from the priority date or could not be elected because they are not bound by Chapter II.

Name and mailing address of the ISA, US
Commissioner of Patents and Trademarks
Box PCT
Washington, DC 20231

Facsimile No. (703) 305-3230

Authorized officer

PAUL A. BELL

Telephone No. (703) 306-3019

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 4521/001PC	<div style="display: flex; justify-content: space-between;"> <div style="text-align: center;">FOR FURTHER ACTION</div> <div style="font-size: small;">see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.</div> </div>	
International application No. PCT/US99/00086	International filing date (day/month/year) 04 JANUARY 1999	(Earliest) Priority Date (day/month/year) 06 JANUARY 1998
Applicant THE VIDEO MOUSE GROUP		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 4 sheets.

☒ It is also accompanied by a copy of each prior art document cited in this report.

1. ☐ Certain claims were found unsearchable (See Box I).
2. ☐ Unity of invention is lacking (See Box II).
3. ☐ The international application contains disclosure of a nucleotide and/or amino acid sequence listing and the international search was carried out on the basis of the sequence listing

☐ filed with the international application.
☐ furnished by the applicant separately from the international application,

☐ but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.

☐ transcribed by this Authority.
4. With regard to the title, ☒ the text is approved as submitted by the applicant.
☐ the text has been established by this Authority to read as follows:
5. With regard to the abstract,

☐ the text is approved as submitted by the applicant.
☒ the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.
6. The figure of the drawings to be published with the abstract is:

Figure No. 1 ☒ as suggested by the applicant.
☐ because the applicant failed to suggest a figure.
☐ because this figure better characterizes the invention.

☐ None of the figures.

Box III TEXT OF THE ABSTRACT (Continuation of item 5 of the first sheet)

The technical features mentioned in the abstract do not include a reference sign between parentheses (PCT Rule 8.1(d)).

NEW ABSTRACT

A human motion following controller(10) is provided by the invention to augment motion of items (e.g., computer cursor or scene view) shown on a computer display. The display (14) is coupled to the computer(20) which controls positioning of the items through operating system controls. A camera(16) captures frames of data corresponding to a first image of at least part of a user (e.g., eyes, hands) at the computer display. Signal processing electronics(18) coupled to the camera (a) detects differences between successive frames of data corresponding to motion of the first image, and (b) communicates differences information to the computer to reposition display of the items through the operating system controls. The items are thus repositioned on the display by an amount corresponding to the motion of first image.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/00086

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : G09G 5/00

US CL : 345/156

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 345/156, 157, 158, 112, 419, 427; 348/154, 155, 169, 170, 171, 172, 77, 78; 382/107; 364/188, 190; 386/117, 702/150; 340/825.19

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS search terms: cursor, camera, motion, movement, disability, handicapped, head, face, hand, eye and body

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4,950,069 A (HUTCHINSON) 21 August 1990, abstract, figures 1-16, column 1 and 2.	1
Y	US 5,168,531 A (SIGEL) 01 December 1992, abstract, figures 1-2, column 1 and 2.	1-79
A	US 5,252,950 A (SAUNDERS ET AL.) 12 October 1993, abstract, figure 1, column 1 and 2.	1
A	US 5,287,473 A (DEERING) 15 February 1994, abstract, figure 1, 6, 7a and 7b, column 1 and 2.	1
Y	US 5,367,315 A (PAN) 22 November 1994, abstract, figure 1, 2 and 4, column 1 and 2.	1-79

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" Document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" Document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"Z" Document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason as specified	"G"	document member of the same patent family
"A" Document referred to in the disclosure, use, exhibition or other means		
"P" Document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

03 JUNE 1999

Date of mailing of the international search report

07 JUL 1999

Name and mailing address of the ISA, US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

PAUL A. BELL

Telephone No. (703) 306-3019

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US99/00086

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,581,276 A (CIPOLLA ET AL.) 03 December 1996, abstract, figure 1, 2, 10 and 16, column 2, lines 20-65	1-79
Y	US 5,617,312 A (IURA ET AL.) 01 April 1997, abstract, figure 1, 2, 4, 5 and 6-11, column 1 and 2.	1-79

20 PR+G

526 Rec'd PCT/PTC 30 JUN 2000

HUMAN MOTION FOLLOWING COMPUTER MOUSE AND GAME CONTROLLER

Related Applications

5 This application is a continuing application of commonly-owned and co-
pending U.S. provisional application number 60/070,512, filed on January 6, 1998
and herein incorporated by reference.

Background

10

The primary human interfaces to today's computer are the keyboard, to enter
textual information, and the mouse, to provide control over graphical information.
These interfaces help users with word processing, presentation software, computer
aided design packages, spreadsheet analyses, and other applications. These
15 interfaces are also widely used for computer gaming entertainment; though they are
often augmented or replaced by a joystick.

In daily use of business software applications, control of cursor position on
the screen requires that the user remove his/her hand from the keyboard in order to
20 use the standard mechanical mouse. The use of the mouse introduces several issues.
In a desk environment, the mouse requires maintenance of space on the desk area.
The mouse cord must also remain free from obstruction to facilitate movement.
Additionally, the use of the mouse is a major contributing factor of carpal-tunnel
syndrome. It would be advantageous therefore to find an alternative to the
25 mechanical mouse.

In computer gaming, game complexity generally requires control of the (i)
mouse and keyboard, or (ii) joystick and keyboard. Further, gaming applications
usually require control in several axes of motion, including forward motion, reverse
30 motion, left turn, right turn, left strafe (slide), right strafe, upward motion,
downward motion. To further complicate game maneuvers and control, many

games permit viewing (within the game environment) in directions different from that in which the vehicle (e.g., the car, or person, simulated within the game) is moving, including up, down, left and right. These many complexities of motion in fact increase or modify the complexity and enjoyment of the game.

5

Nevertheless, these complexities require that the user have utmost dexterity and control of his/her body. One object of the invention, therefore, is to offer alternative approaches to human-computer interfaces for those incapable of using standard devices (e.g., mouse, keyboard and joystick) such as due to disability.

10

Another object of the invention is to provide an alternative input device for laptop computers. Laptop computers are used in locations which do not allow the use of a mouse, in airplanes or during business meetings in which there is no room to operate the mouse. Through the use of either a clip on camera or a camera built into the laptop display, the laptop user can control the mouse position or use the camera for teleconferencing while on the road.

15

Other objects of the invention are to replace or augment existing human computer interfaces to facilitate enhanced gaming and/or control within game environments.

20

In the prior art, certain systems exist which attempt to reduce the amount of physical interaction required with game controllers. However, such systems are prohibitively expensive to the general public as their costs are driven by techniques and algorithms which detect user head motion based upon a detectable target worn by the user. Other costly and cumbersome systems require the user to wear apparatus which emits or detects a signal. It is thus one other object of this invention to provide a system which detects user motion without the aid or augmentation of artificial devices placed on the user operator.

25

Another object of the invention is to provide a means of human control of a graphical computer interface through the physical motion of the user in order to

30

control the activity of a cursor in the manner usually accomplished with a computer mouse.

5 A further object of the invention is to provide additional degrees of freedom in the human computer interface in support of computer games and entertainment software.

Yet another object of the invention is to provide dual use of teleconferencing and video electronics with gaming and computer control systems.

10

These and other objects will be apparent in the description which follows.

Summary of Invention

15 As used herein, "cursor" means a computer cursor associated with a computer screen. "Scene view" means the view presented on a computer display to a user. For example, one scene view corresponds to the scene presented to a user during a computer game at any given moment in time. The game might include displaying a scene whereby the user appears to be walking in a forest, and through
20 trees. In another example, a cursor might also be visible in the scene view as a mechanism for the user to select certain events or items on the scene (e.g., to open a door in a game, or to open a folder to access computer files).

As used herein, "camera" refers to a solid state instrument used in imaging.
25 Typically, the camera also includes optical elements which refract light to form an image on the camera's detector elements (typically CCD or CMOS). For example, one camera of the invention derives from a video-conferencing camera used in conjunction with Internet communication.

30 In one aspect, the invention provides systems and methods to control computer cursor position (or, for example, the scene view or game position as

displayed on the computer display) by motion of the user at the computer. A camera rests on or near to the computer, or built into the computer, and connects therewith to collect "frames" of data corresponding to images of the user. These images provide information about user motion, over time. Software within the computer assesses these frames and algorithmically adjusts cursor motion (or scene view, or mouse button, or some other operation of the computer) based upon this motion. The motion may be imparted by up-down or left-right motion of the user's head, by the user's hands, or by other motions presented to the video camera (such as discussed herein). In one aspect, a close up view of the users facial features is used to impart a translation in the cursor (or scene view) even through the features in fact rotate with the user's head. In yet another aspect, the rotation is used to generate a corresponding rotation in computer game scene imagery.

In one aspect, the invention also provides a human factors approach to cursor movement in which the user's rate of motion determines the relative motion of the cursor (or scene view). By way of example, the faster the user's head travels over a set distance, the further the corresponding cursor movement over the same time period.

In other aspects of the invention, the camera is either (a) a visible light camera utilizing ambient lighting conditions or (b) a camera sensitive in another band such as the near infrared ("IR"), the IR, or the ultraviolet ("UV") spectrum. In the latter case (b), the illumination preferably emanates from a source such as an IR lamp which is beyond human sensory perception. The sensor is typically mounted facing the user so as to capture a picture of the user's face in the associated electromagnetic spectrum. The lamp is typically integrated with the camera housing so as to facilitate production and ease of consumer set-up.

In one aspect, a system of the invention provides an IR camera (i.e., a camera which images infrared radiation) to image the user's face and to gauge the user's stress level associated with a game on the computer. As the user's intensity increases

(such as in a fast moving computer game using a joystick or the methods discussed herein), the system detects increased heat intensity on the user's face, forehead or other body part by the imagery of the IR camera. This information is fed back into the game processor to provide further enhancement to the game. In this manner, the system gauges the user's reaction to the game and modifies game speed or operation in a meaningful way. For example, suppose such a system determined that a particular user was bored of the present game speed (a determination of boring can be made by assessing low IR output over large portions of the user's face). The computer processor and game software can then cooperate to increase the gaming speed and thereby increase this particular user's stress. Games of the invention are thus made and sold to users with varying intelligence, age and/or computer familiarity; and yet the system always "pushes the envelope" for any given user so as to make the game as interesting as possible, automatically.

In accord with one aspect of the invention, images captured by the sensor are processed by a digital signal processor ("DSP") located either (a) in a PC card within the host computer or (b) in a housing integrated with the sensor. In case (a), sensor frames are sent to the PC card; and detected user motion (sometimes denoted herein as "difference information") is communicated to the user's operating system via a PCI (or USB or later standard) bus interface. These difference information commands are interpreted by a low overhead program resident at the user's main processor, which either updates the cursor position on the screen or provides motion information to the user's computer game (e.g., so as to change the scene view). In case (b), the DSP is contained within the camera housing; and frames are processed local to the camera to determine difference information. This information is then transmitted to the computer by a cable that connects to a bus port of the computer so that the host processor can make appropriate movements of the cursor or scene view. In another aspect, the DSP is mounted in the camera housing such that the camera/signal processing subsystem produces signals which emulate the mouse via the mouse input connector.

In an alternative configuration, frames of image data are sent directly to the host computer through the computer bus; and that image data is manipulated by the computer processor directly. With increasing computer processing speed, it is expected that sensor data frames can be sent directly to the host processor for all processing needs, in which case the PC card and/or separate DSP are not required. Although this is possible today, the update rates are likely too slow for practicality. Once GHz processors are on the market, a separate DSP may no longer be needed.

In one aspect of the invention, pixel format or pixel density of the camera drives the accuracy of the system. Higher pixel density in the image of the user's face, for example, increases the attainable resolution and cursor control (or the attainable control of scene view motion). Camera formats of 240 vertical by 320 horizontal generally provide satisfactory performance. The number of pixels that may be utilized is determined by system cost factors. Greater numbers of pixels require more powerful DSPs (and thus more costly DSPs) in order to process the image sequences in real time. Current technology limits the processing density to a 64x64 window for consumer electronics. As prices are reduced, and power increases, the densities can increase to 128x128, 256x256 and so on. While 64x64 density is satisfactory for general household users, a higher fidelity system using a greater number of pixels is possible, in accord with the invention, for higher end applications at a proportionally higher cost.

In one aspect, the data transfer rate from the camera is 30 frames/second at 240x320 pixels per frame. Assuming eight bits per pixel, the digital data transfer rate is therefore 18.432 megabits/second. This is a fairly high transfer rate for consumer products using current technology. While the data transfer can be either analog or digital, the preferred method of image data transfer for this aspect is via a standard RS170 analog video interface.

In accord with one aspect, a system of the invention defines two imaging zones (either within a single camera CCD or within multiple CCD cameras housed

within a single housing). One imaging zone covers the user's head; and the other covers the user's eyes. This aspect includes processing means to process both zones whereby movement of the user's head provides one mechanism to control cursor movement (or scene view motion), and whereby the user's eyes provide another
5 mechanism to control the movement. In essence, this aspect increases the degrees of freedom in the control decision making of the system. By way of example, a user might look left or right within a game without moving his head; but by assessing movement of the user's eyes (or the pupils of those eyes), the scene view can be made to rotate or translate in the manner desired by the user. Further, a user might
10 move his head for other reasons, and yet not move her eyes from a generally forward looking position; and this aspect can assess both movements (head and eyes) to select the most appropriate movement of the cursor or scene view, if any.

In another aspect, a system of the invention utilizes a camera with zoom
15 optics to define the user's pupil and to make cursor or scene views move according to the pupil. In another aspect, the system incorporates a neural net to "learn" about a user's eye movements so that more accurate movements are made, over time, in response to the user's eye movement.

20 In still another aspect, a neural net is used to learn about other movements of the user to better specify cursor or scene view movement over time.

In yet another aspect of the invention, a system is provided with two CCD arrays (either within a single camera body or within two cameras). The arrays
25 connect with the user's computer by the techniques discussed herein. One CCD array is used to image the user's head; and the other is used to image the user's body. Motion of the user is then evaluated for both head and body movement; and cursor or scene view movement is adjusted based upon both inputs.

30 In another aspect of the invention, a single CCD is used to image the user. However, alternate frames are zoomed, electronically, so that one frame views the

user's head, and the next frame views the user's eyes. With the algorithm discussed herein, these separate frame sequences (one for the eyes, one for the head) are processed separately and evaluated, together, to make the most appropriate cursor or scene view movement. If for example, the system clocks at 30Hz, then one set of
5 frame sequences operates at 15Hz, and the other at 15Hz. However, the advantage is that two movement information sets can be evaluated to invoke an appropriate movement in the cursor or scene view.

Those skilled in the art should appreciate that different frame rates can be
10 used; and frame rates for either sequence (head or eyes) can occur at different rates too. Further, the separate frame sequences can utilize other body parts, e.g., the head and the hand, to have two movement evaluations. Alternatively, a separate camera (or CCD array) can be used to image other body parts, for example one camera for the head and one for the hand.

15 The invention also provides methods for shifting cursor or scene views in response to user movement. In one aspect, the scene view shifts left or right when the user shifts left and right. In another aspect, the scene view rotates when the user's head rotates. This last aspect can be modified so that such rotation occurs so
20 long as the eyes do not also rotate (in this situation, the user's head rotates, indicating that she wishes the scene view to rotate; but the eyes do not, indicating that she still watches the game in play). In another aspect, the scene view rotates in response to the user's hand rotation (i.e., a camera or at least a CCD array of the system is arranged to view the player's hand).

25 In another aspect, the invention provides a multi-zone player gaming system whereby the user of a particular computer game can select which zone operates to move the cursor or the scene view. By way of example, the system can include one zone corresponding to a view of the user's head, where frames of data are captured
30 by the system by a camera. Another zone optionally corresponds to the user's hand. Another zone optionally corresponds to the user's eyes. Each zone is covered by a

camera, or by a CCD array coupled within the same housing, or by optical zoom zones within a single CCD, or by separate optical elements that image different portions of the CCD array. By way of example, two zones can be covered with a single CCD array (i.e., a camera) when the zones are the user's head and eyes. The camera images the head, for one zone, and images the eyes in another zone, since the zones are optically aligned (or nearly so). However, two cameras (or optionally two CCD arrays with separate optics) can view two zones such as the user's head and the user's hand. Combinations of zones is also possible and envisioned in accord with the invention.

Zones in a single camera can also be identified by the computer by prompting the user for motion from corresponding body parts. For instance, the computer identifies the head zone by prompting the user to move his head. Then the computer identifies the foot zone by having the user move his foot. Once the zones are identified, the motion of each of these individual zones is tracked by the computer and the regions of interest in the camera image related to the zones moved as the targets in the zones move with respect to the camera.

In one aspect, the invention provides a system, including a camera and edge detection processing subsystem, which isolates edges of the user's body, for example, the side of the head. These edges are used to move the cursor or scene view. For example, if the left edge of head is imaged onto column X of one frame of the CCD within the camera, and yet the edge falls in column Y in the next frame, then a corresponding movement of the cursor or scene view is commanded by the system. For example, movement of the edge from one column to the next might correspond to ten screen pixels, or other magnification. In one aspect, this magnification is selected by the user. Up and down motion can also be detected by similar edge detection. For example, by imaging the user's chin, an edge movement in the up or down dimension is formed (e.g., if the bottom edge of the chin moves from one row to the next, in adjacent frames, then a corresponding movement of the cursor or scene view is made - magnification again preferably set manually with a

default starting magnification). Other images can also serve to define edges. For example, in one aspect, a user's eyelash can be used to move the cursor (or scene view) up or downwards; though typically the eye blink is used to reset the cursor command cycle.

5

In one aspect, an optical matched filter is used to center image zones onto the appropriate images. For example, as discussed above, one aspect preferably utilizes 64x64 pixels as the image frame from which cursor motion is determined. Many cameras have, however, many more pixels. These 64x64 arrays are therefore preferably established through matched filtering. By way of example, an image of a standard pair of user's eyes is stored within memory (according to one aspect of the invention). This image field is cross-correlated with frames of data from the actual image from the camera to "center" the image at the desired point. With eyes, specifically, ideally the 64x64 sample array is centered so as to view both eyes within the 64x64 array. Similarly, to process sequences of head data, a standard head image is stored within memory, according to one aspect, and correlated with the actual image to center the head view.

Those skilled in the art should appreciate that an appropriate frame size can be established from an image having more or fewer pixels, by redundantly allocating data into adjacent pixels or by eliminating intermediate pixels, or similar technique.

In another aspect, a camera is provided which optically "zooms" to provide optimal imaging for a desired image zone. By way of example, the invention of one aspect takes an image of the user's head, determines the location of the user's eyes (such as by matched filtering), and optically zooms the image through movement of optics to provide an image of the eyes in the desired processing size format.

Many aspects of the invention are preferably enhanced by autofocus. Specifically, it is often desirable to have a crisp image of the user (or a part of the user, e.g., the user's eyes) in order to accurately process desired cursor or scene view

movement. Thus, autofocus capability preferably operates in most of the aspects of the invention where imaging is a feature of the processing.

In one aspect, the camera utilizes a very small aperture which results in a very large depth of field. In such a situation, autofocus is not required or desired. The
5 optical requirements for the lenses are also reduced.

The invention thus provides several advantages over the art. For example, game controllers can now include feedback corresponding to the user's actual movement. By way of another example, if the user moves left or right (or head or
10 hand or eyes move left or right, depending on the image zone), then the cursor (or scene view) can also be set to move left or right. When the user twists her head, for example, the scene view can also be made to rotate, reflecting that movement.

Those skilled in the art should appreciate that the direction in which the scene
15 moves, left or right, is a matter of design choice. That is, certain games might find it desirable to move the opposite direction from what the user moves, to add certain challenges to the game. Further, in other aspects, this direction can change during the game to further complicate game control.

20 In accord with one aspect of the invention, a processing subsystem (connected with the camera) is used to make cursor movement (or scene view movement) correspond to user's motion. This processing subsystem of another aspect further detects when the user twists his head, to add an additional dimension to the movement.

25

In one aspect, a system of the invention includes an IR detector which is used to determine when a person sweats or heats up (by imaging, for example, part of the user's head onto the IR detector); and then the system adjusts game speed in a way corresponding to this movement. Alternatively, a heartbeat sensor is tied to the
30 person to sense increased excitement during a game and the system speeds or slows the game in a similar manner. Note that a heartbeat sensor can be constructed, in one

aspect of the invention, by thermal imaging of the user's face, detecting blood flow oscillations indicative of heartbeat. In other aspects, the heartbeat sensor is physically tied to the user, such as within the computer mouse or joystick.

5 In one aspect, a computer of the invention adapts to user control as selected by a particular user. For example, in the case of a handicapped person, a particular user might select certain hand-movements, e.g., a single finger up, to move the cursor up; and another finger down to move the cursor left. An infinite combination of controls can be established; however this is one advantage of the invention in that
10 users with many different disabilities can program cursor or scene view movement. In one aspect, a neural network is used to assist the processing system in establishing proper cursor movement. In another aspect, the computer for example learns to print something by movement of the user's finger (or other body part).

15 In one aspect, tipping of the user's head (or other body part, or object) is used to provide another degree of freedom in moving the cursor or adjusting the scene view. By way of example, a tilt of the head, as imaged by the camera, can be set to command a rotation of the scene view.

20 In still another aspect, a camera of the invention uses autozoom to move in and out of a given scene view. By way of example, the camera is first focussed on the user's face in one frame; but in a subsequent frame the camera must focus to closer to compensate for the fact that the user moved closer to the camera (typically, the camera is on the monitor, so this also means that the user moved closer to the scene
25 view). This autofzoom is used, in one aspect, to make the scene view appear as if the user is "creeping" into the scene. By moving the scene in and out, the user will perceive that he is moving in or out of the scene view.

30 In another aspect, a camera images an object held by the user. Preferably, the object has a well-defined shape. The system images the object and determines difference information corresponding to movement of the object. By way of example,

rotating the object upside down results in difference information that is upside down; and then the scene view inverts by operation of the system. In another example, twisting of the object rotates the scene view left or right, or rotates the scene in the direction of the twisting.

5

In another aspect, two cameras image the user: one camera pointed at the front of the users face or hand and the other down at the top of the users head or hand. The front facing camera is used to detect rotational and linear translation in up-down and left-right directions. The top viewing camera determined front-back, left right translation. The front-back translation observed by the top camera is used to control forward and back motion in the users 3-D view. The top sensed left-right translation controls the users left right slide or strafe. The top sensed left-right motion is removed from the front view left-right translation with the remaining front view measure representative of left-right twist. All of the front view up-down translation can be interpreted as up-down twist.

Brief Description of the Drawings

Figure 1 illustrates one human computer interface system constructed according to the invention;

Figure 1A shows an exemplary computer display illustrating cursor movement made through the system of Figure 1;

Figure 1B illustrates overlayed scene views, displayed in two moments of time on the display in Figure 1, of a shifting scene made in response to user movement captured by the camera of Figure 1;

Figure 1C shows an illustrative frame of data taken by the system of Figure 1;

Figure 2 illustrates selected functions for a printed circuit card used in the system of Figure 1;

Figure 3 illustrates an algorithm block diagram that preferably operates with
5 the system of Figure 1;

Figure 4 illustrates one preferred algorithm process used in accord with the invention to determine and quantify body motion;

10 Figure 5 shows one process of the invention for communicating body motion data to a host processor, in accord with the invention, for augmented control of cursor position or scene view;

Figure 5A shows representative frame of data of a user taken by a camera of
15 the invention, and further illustrates adding symbols to key body parts to facilitate processing;

Figure 6 illustrates a two camera imaging system for implementing the teachings of the invention;

20 Figure 7 illustrates two positions of a user as captured by a camera of the invention; and Figure 7A illustrates two positions of a scene view on a display as repositioned in response to movement of the user illustrated in Figure 7;

25 Figure 8 illustrates motion of a user - and specifically twisting of the user's head - as captured by a system of the invention; Figure 8A illustrates a first scene view corresponding to a representative computer display before the twisting; Figure 8B illustrates a second scene view corresponding to a rotation of the first scene view in response to the twisting by the user; Figure 8C shows processing features of the
30 processing section of Figure 8; and Figure 8D illustrates multiple image frames

stored in memory for matched filtering with raw images acquired by the system of Figure 8;

Figure 9 illustrates a two camera system of the invention for collecting N
5 zones of user movement and for repositioning the cursor or scene view as a function of the N movements; Figure 9A illustrates a representative thermal image captured by the system of Figure 9; and Figure 9B illustrates process methodology for processing thermal images as a real time input to game processing speed, in accord with the invention;

10 Figure 10 illustrates another two camera system of the invention for targeting multiple image movement zones on a user, and further illustrating optional DSP processing at the camera section;

15 Figure 11 illustrates framing multiple movement zones with a single imaging array, in accord with the invention;

Figure 12 illustrates framing a user's eyes in accord with the invention; and Figure 12A shows a representative image frame of a user's eyes;

20 Figure 13 illustrates one system of the invention, including zoom, neural nets, and autofocus to facilitate image capture;

Figures 14, 14A and 14B illustrate autofocus motion control in accord with
25 the invention;

Figure 15 illustrates one other motion detect system algorithm utilizing edge detection, in accord with the invention;

30 Figure 16 illustrates one other motion detect system algorithm utilizing well-characterized object manipulations , in accord with the invention;

Figure 17 illustrates one other motion detect system algorithm utilizing varied body motions, in accord with the invention;

5 Figure 18 illustrates a two camera system of the invention with a camera observing the user's face while the other observes the top of the user's head; and

Figure 19 shows a blink detect system of the invention.

10 Detailed Description of the Drawings

Figure 1 illustrates, in a top view, certain major components of a human computer interface system 10 of the invention. A user 12 of the system 10 sits facing a computer monitor 14 with display 14a. A camera 16 is mounted on the computer
15 monitor 14 facing the user 12. In the illustrated embodiment, the camera 16 is mounted in such a way that the user's face 12a is imaged within the camera's field of view 16a. However, as discussed herein, the camera 16 can alternatively image other locations, such as the user's hand, eyes, or on other objects; so imaging of the user's face, in Figure 1, must be considered illustrative, rather than limiting. Further, the
20 camera location can also reside at places other than on top of the monitor 14.

With further regard to Figure 1, the camera 16 interfaces with a printed circuit card 18 mounted within the user's computer chassis 20 (which connects with the monitor 14 by common cabling 20a). The camera 16 interfaces to the printed circuit
25 card 18 via a camera interface cable 22. The circuit card 18 also has processing section 18a, such as a digital signal processing ("DSP") chip and software, to process images from the camera 16.

In operation, the camera 16 and card 18 capture frames of image data
30 corresponding to user movement 25. The processing section 18a algorithmically processes the image data to quantify that movement 25; and then communicates this

information to the host processor 30 within the computer 20. The host processor 30 then commands movement of the computer cursor in a corresponding movement 25a, Figure 1A (Figure 1A illustrates a representative front view of the display 14a, and also illustrates movement 25a of the cursor 26 moving within the display 14a in response to user movement 25).

Figure 1B illustrates an alternative (or supplemental) process whereby the scene view shifts in response to user movement 25. Specifically, Figure 1B illustrates a first scene view 35a, which generally corresponds to a forest prior to the user's movement 25; and an overlaid scene view 35b (shown in dotted line, for purposes of illustration) that is shifted by an amount 37 in response to the user's movement 25. The shift 37 in the scene view 35 is accomplished by combined operation and processing of the processing section 18a and host CPU 30.

Figure 1C shows a representative frame 41 of data 43 as taken by the camera 16. As illustrated, data 43 represents the user's face 12a taken at a given moment of time. Subsequent frames (not shown) are used to determine user motion 25 relative to the frame 41, as discussed herein. The frame 41 is made up of the plurality of pixel data 45, as known in the art.

20

Figure 2 illustrates certain functions processed within the printed circuit board 18 of Figure 1. A camera interface circuit 50 receives video data from the camera 16 through interface cable 22. This data can be RS170 format. Circuit 50 decodes the analog video data to determine video timing signals embedded in the analog data. These timing signals are used for control of the analog-to-digital (A/D) converter included in circuit 50 that converts analog pixel data into digital images. In the preferred embodiment, the analog data is digitized into 6-bits, though any number of bits greater may be acceptable and/or required for features as discussed herein.

30

The frame difference electronics 52 receives digital data from the camera interface circuit 50. The frame difference electronics 52 include a multiple frame memory, a subtraction circuit and a state machine controller/memory addresser to control data flow. The frame memory holds previously digitized frame data. As each digitized pixel is received by the frame difference electronics 52, the corresponding pixel from a previous frame is read from the frame memory and subtracted from the current frame. The preferred implementation uses the frame just previous to the current frame, though an older frame which resides in the frame memory may be used. The resulting difference is output to an N-frame video memory 54. The new frame pixel data is then stored into the frame memory of the frame difference electronics 52.

The N frame video memory electronics 54 either receives differenced frames output by the frame difference electronics 52 (discussed above) or raw digitized frames from the camera interface 50. The choice of where the data derives from is made by software resident on the DSP 56. The frame video memory 54 is sized to hold more than one full frame of video and up to N number of frames. The number of frames N is to be driven by hardware and software design.

In the preferred embodiment, the DSP 56 implements an algorithm discussed below. This algorithm determines the rate of head motion of the user in two dimensions. The digital signal processor 56 also detects the eye blink of the user in order to emulate the click and double click action of a standard mouse button. In support of these functions, the DSP 56 commands the N frame video memory 54 to supply either the differenced frames or the raw digitized frames. The digital signal processor thus preferably utilizes a supporting program memory 58 made up of electrically reprogrammable memory (EPROM) and data memory 59 including standard volatile random access memory (RAM). The DSP 56 also interfaces to the PCI bus interface electronics 60 through which cursor and button emulation is passed to the user's main processor (e.g., the CPU 30, Figure 1). The PCI interface 60 also passes raw digitized video to the main processor as an optional feature.

Interface 60 also permits reprogramming of program memory 58, to allow for future software upgrades permitting additional features and performance.

The PCI interface electronics 60 thus provides an industry standard bus
5 interface supporting the aforementioned communication path between the printed circuit card 18 and the user's main processor 30.

With optional MPEG compression electronics 62, the printed circuit card 18 and camera 16 can provide compressed video to the user's main processor 30. This
10 compressed video supports using the system 10 in teleconferencing applications, providing dual use as either human computer interface system 10 and/or as a teleconferencing system in an economical solution to two distinct applications.

Figure 3 describes one preferred head motion block diagram algorithm 70
15 used in accord with the invention. Not all of the functions shown in Figure 3 are implemented in software in the DSP 56. Rather, this algorithm relies on the correlation of images from one frame to the next, and particularly relies on the use of frame differenced images in the correlation process. The frame differencing operation removes parts of the camera images that are unchanged from the previous
20 frame. For example, room background (such as object 13, Figure 1) behind the user 12 is removed from the image. This greatly simplifies detection of feature motion. Even the image of the user's face image consists of regions of uniform illumination such that even with the user's facial motion, these uniform regions (i.e. cheeks, forehead, chin) may also be removed. The user's face 12a also consists of typically
25 dynamic features such as the nose, eyes, eyebrows and mouth, each of which typically has enough spatial detail that will be evident in the differenced image. As the user moves his face with respect to room lighting, the shape and distribution of these features will change; but the frame rate of the camera 16 ensures that these features look similar from one frame to the next. The correlation process therefore
30 operates to determine how these differenced features are moving from one frame to the next in order to determine user head motion 25.

The algorithm of block diagram 70, Figure 3, receives video images 72 of the user as imaged by camera 16 over time. Each received image is passed to both a frame memory 74 and a differencer 76. Though the preferred embodiment is to buffer a single frame in memory 74, the memory 74 may optionally store many frames, buffered such that the first frame input is the first frame output (FIFO). The delayed frame is read from the frame memory 74 and subtracted from the current frame using the differencer 76. Frame output from the differencer 76 is provided to both a correlation process 78 and a difference frame memory 80.

Like the frame memory 74, the preferred embodiment of frame memory 80 utilizes a single difference frame; however the difference frame memory 80 can hold many difference frames in sequence in a FIFO arrangement. The delayed difference frame is read from the difference frame memory and provided to the correlation function 78. The correlation process 78 determines the best combination of row and column shifts in order to minimize the difference between the current difference frame and the delayed difference frame. The number of rows and columns required to align these difference images provides information as to the user's motion.

The best-fit function algorithm 82 determines the row and column shift to provide optimal alignment. In the case of a classical correlation process, the best-fit function can consist of a peak detect algorithm. This algorithm may either be implemented in hardware or in software.

The best-fit function algorithm determines relative motion in rows and columns of the observed user's features. The cursor update compute function algorithm 84 translates this measured motion into the position change required of the cursor (e.g., the cursor 26, Figure 1A). Typically, this is a non-linear process that, with greater head motion, the cursor moves a non-proportionally greater distance. For example a 1-pixel user motion can cause the cursor to move one screen pixel while a 10-pixel user motion may cause a 100-pixel screen cursor motion. However,

these magnifications can be adjusted for desired result. This algorithm may either be implemented in hardware or in software such as through an ASIC or FPGA.

Video cursor control 86 provides a user interface to enable and disable the operation of cursor control described above. This control is implemented, for example, through a combination of keystrokes on the user's keyboard (for example as connected to the host computer 20, Figure 1). Alternatively, cursor control is activated or deactivated by sensing the eye-blink of the user (or some other predetermined movement). In this alternative embodiment, an output signal 85 from the correlation section 78 is sent to the video enable section 86; and the output signal 85 corresponds to blink data from the user's face 12a (Figure 1A). In another embodiment, the video cursor control section 86 activates or deactivates cursor control by recognizing voice commands. A microphone 87 detects the user's voice and a voice recognition section 89 converts the voice to certain activate or deactivate signals. For example, the section 89 can be set to respond to "activate" as a voice command that will enable cursor control; and "deactivate" as a command that disables cursor control.

The functionality of the video cursor control 86 provides the user with the equivalent of a mouse pick-up, put-down action. As the user moves the cursor from left to right across the screen, the user would de-activate motion-based cursor control in order to allow the user to move his head back to the left. Once the user has recentered his head, the user would once again activate the cursor control and continue to move the cursor about the screen. The activation/deactivation of the mouse input is represented by the switch 90, such that the open position of the switch disables human motion control of the cursor and supplies a zero change input to the summation operation 92 in such conditions.

Those skilled in the art should appreciate that control of scene view may also be implemented by an algorithm such as shown in Figure 3. Specifically, a similar

algorithm can provide movement of the current scene view, in accord with the invention.

With video cursor control enabled, the result of the cursor update compute
5 function 84 is added to the known current cursor position at the summing operation
92. This summation has an x component and a y component. The result of the
summation 92 is used to update the cursor position (or scene view) on the user's
screen via the user's operating system. Cursor position may thus be controlled by
both user motion as well as the motion imparted by another input device such as a
10 standard computer mouse.

Figure 4 provides a detailed description of the preferred implementation of
the algorithm described in functions 74, 76, 78 and 82 of Figure 3. Video data is
received by the processing electronics in both a single frame memory 100 and a
15 differencer 102. The output of the frame memory 100 is also provided to the
differencer 102 such that the previous frame is subtracted from the current frame.
This differenced frame is then processed by a two dimensional FFT 104. The complex
result of the FFT 104 is provided to a complex multiplier 106 and a complex memory
108. The complex memory 108 is the size of the processed image, each location
20 containing both a real and imaginary component of a complex number. With each
new FFT operation 104, the previous FFT result, contained in the complex memory
108, is provided to the conjugate operation 110. The complex conjugate of each
element is computed and provided to the complex multiplier 106. In this manner,
the FFT of the previous frame difference is conjugated and multiplied against the
25 FFT of the current difference image.

The two dimensional array of complex products output by the complex
multiplier 106 is provided to a two dimensional inverse FFT operation 112. This
operation creates an image of the correlation function 114 between the latest pair of
30 difference images. The correlation image is processed by the peak detection function
114 in order to determine the shift required in aligning the two difference images.

The x-y magnitude of this shift is representative of the user's motion. This x-y magnitude is provided to the software used to update the cursor position as described in Figure 3.

5 Figure 5 shows an algorithm process 130 of the invention which applies motion correlation operations over sub-frames of the video image. This allows motions of various body parts to convey input with specialized meaning to applications operating on the host computer. In addition to head motion, motion of the hands, arms and legs provide for greater degrees of freedom for the user to
10 interact with the host application (e.g., a game). Commands of this type are useful in combative games where computer animated opponents fight under control of the user. In that instance, the hand, arm and leg motions of the user become punch, chop and kick commands to the computer after process 130. This command mode can also be used in situations where the user does not have ready access to a keyboard, to
15 augment cursor control of the previously described head position correlator.

Process 130 identifies the functions required to derive commands from general motions of the user's body. The scene analyzer function 132 receives digitized video frames from the camera (e.g., the camera 16 of Figure 1) and
20 identifies sub-frames within the video for tracking various parts of the user's body. The frame difference function 134 and correlator function 136 provide similar functions as processes 74, 76 and 78 of Figure 3. The correlation analyzer 138 receives correlated difference frames from the correlator function 136 and sub-frame definitions from the scene analyzer 132. The correlation analyzer 138 applies a peak
25 detection function to each sub-frame to identify the shift required to achieve best alignment of the two images. Correlation peaks occurring in the center of the sub-frame indicate no motion, while peaks occurring elsewhere indicate the direction and magnitude of the user's motion. The motion interpreter 140 receives motion vectors for each sub-frame from the correlation analyzer 138. The motion interpreter
30 140 links the motion vector from each sub-frame with a particular body segment and passes this information onto the host interface 142. The host interface 142 provides

for communication with the host processor (e.g., CPU 30, Figure 1). It sends data packets to the host to identify detected body motions, their directions and their amplitudes. The host interface 142 also receives instruction from the host as to which body segments to track which it then passed along to the motion interpreter 140 and the scene analyzer 132.

The scene analyzer 132 first identifies the location of the user's body in the image and locates the position of various parts of the user's body such as hands, forearms, head, and legs. The techniques and methods used to identify the user's body location and body part positions can be accomplished using techniques well known to those skilled in the art (by way of example, via matched filtering). Body identification can also be augmented by marking different locations on the user's body with unique visual symbols. Unique symbols are assigned to key body joints such as elbows, shoulders, hands, neck, knees, and waist and are mounted on the body. See for example Figure 5A.

Figure 5A illustrates one frame 149 of data of an image of the user 150 as taken by a camera of the invention. The image corresponds to a full body image of the user 150, including arms 151, legs 152, elbows 151a, hands 153, head 154, neck 155, ears 156, and forehead 157. These parts 151-157 are identified by processes of the invention (e.g., spatial location in the image, by matched filtering or other image recognition technique), and the image is preferably marked with unique symbols (e.g., "X" for center of the face, "Y" for center of the hand 153, "T" for center of the user's foot, "Z" for body center, and "F" for forehead 157).

25

With further reference to Figure 5, process 130 locates various body parts and preferably marks them with symbols to fill in connecting logic (e.g., the left wrist and left elbow symbol identify the location of the left forearm). Once the user's body parts are located, sub-frames surrounding each of the body segments identified by the host processor are generated. A sub-frame is a generally regularly shaped region within the image that surrounds a particular body part. The sub-frames are sized to

30

center the subject body part in the sub-frame and to provide enough room around the body part to accommodate typical body motions. One sub-frame 160 is shown in frame 149, Figure 5A, surrounding the user's foot "T". The scene analyzer 132 will generally not operate on each frame of video since continuously changing the sub-frames adds unnecessary complication to the correlation analyzer 138. Instead, the scene analyzer 132 runs as a background process updating the sub-frame locations periodically.

Figure 4 provides a detailed description of one algorithm which can be used to implement processes 134-138 of Figure 5.

The invention of one embodiment can thus track the motion of the user's body using symbols attached to key joints. As an example, the position of the user's left lower arm can be determined by locating the unique symbol for the left hand "Y1" and left elbow "●". Unique symbols thus allow the processor to rapidly locate each portion of the user's body in a video frame. To determine the motion of a particular part of the user's body, the algorithm (e.g., Figure 4) compares the position of the relevant body parts in consecutive frames and determines how they moved (for example, using geometry). Once motion is determined, it is then passed to the host CPU where the motion is acted on as appropriate for the particular application.

Figure 6 illustrates a two camera system 200, constructed according to the invention. The cameras 202a, 202b are arranged to view separate parts of the user: camera 202a images the user's face 204; and camera 202b images the user's hand 206. The cameras 202 conveniently rest on top of the computer display 208 coupled to the host computer 210 by cabling 216. The cameras 202 couple to the signal processing card 212 residing within the computer 210 by cabling 213. As discussed herein, motion of the user's head 204 and/or hand 206 are detected by the signal processing card 212, and difference information is communicated to the computer's CPU 210a

via the computer bus 214. This difference information corresponds to composite movement of the head 204 and hand 206; and is used by the CPU 210a to command movement of display items on the display 208 (for example, the display items can include the cursor or scene view as shown on the display 208 to the user).

5 Information shown on the display 208 is communicated from the computer 210 to the display 208 along standard cable 216.

Figures 7 and 7A illustrate how motion of a user's head is for example translated to motion of the cursor and/or scene view, in accord with the invention.

10 Figure 7 shows a representative image 220 of a user captured within a frame of data by a camera of the invention. Figure 7 also shows a representative image 222 (in dotted outline, for clarity of illustration) of the user in a subsequent frame of data, indicating that the user moved "M" inches. Figure 7A illustrates corresponding scene views on a computer display 224 that is coupled to processing algorithms of

15 the invention (i.e., within a system that includes a camera that captures the images 220, 222 of Figure 7). The display 224 illustratively shows a scene view that includes a road 224a that extends off into the distance, and a house 224b adjacent to the road 224a. A computer cursor 224c is also illustratively shown on the display 224 as such a cursor is common even within computer games, providing a place for the user to

20 select items (such as the road or house 224a, 224b) within the display 224. The display 224 also shows, with dotted outlines 226, the scene view of road and house which are shown on the display 224 after motion by the user from 220 to 222, Figure 7 (the cursor 224c is for example repositioned to position 224c'). The repositioning of the scene view from 224a, 224b to 226 occurs immediately (typically much less than

25 one second) after the movement of the user of Figure 7 from 220 to 222. The scene view is repositioned by x-pixels on the display 224, so that M/x corresponds to the magnification between user movement and scene view repositioning. This magnification can be set by parameters within the system; and can also be set by the user, if desired, at the computer keyboard. Furthermore, the rate at which the scene

30 view moves the distance of x-pixels preferably occurs at the same rate as the rate of travel along distance M. Alternatively, the magnification can be dependent on the

rate of motion such that a larger displacement of x-pixels will occur for a given motion M if the rate of change of M is larger.

Figure 8 illustrates a further motion that can be captured by a camera of the invention and processed to reposition a scene view, as shown in Figures 8A and 8B. More particularly, Figure 8 illustrates a camera 250 connected to a processing section 252 which converts user motion 254 to corresponding repositioning of the computer scene view. As above, the user 256 is captured by the camera's field of view 258 and frames of data are captured by the processing section 256. In Figure 8, motion 254 corresponds to a twisting of the user's head 256; and processing section 252 detects this twisting and provides repositioning information to the host computer (not shown). Figure 8A shows a representative scene view 260 on a display 262 coupled to the host computer. Figure 8B illustrates repositioning of the scene view 260' after the processing section 252 detects motion 254 and updates the host computer with difference information (e.g., that information which the host computer uses to rotate or translate the scene view).

Figure 8A also illustrates the intent of the rotating scene view feature. In Figure 8A, a person 260a is shown in the scene view 260, except that the person 260a is almost completely obscured by the edge 262a of the display 262. By twisting the head 256 in motion 254, the scene view 262 is rotated in the corresponding direction - as shown by scene view 260' in Figure 8B - so that the user 260a' is completely visible within the scene view 260'.

Figure 8C illustrates further detail of the processing section 252. Camera data such as frames of images of a user are input to the section 252 at data port 266. The data are conditioned in the image conditioning section 268 (for example, to reduce correlated noise or other image artifacts). Thereafter, the camera data is compared and correlated in the image correlation section 270, which compares the present frame image with a series of stored images from the image memory 272. In the preferred embodiment, the present data image frame 249 is cross-correlated with

each of the images within the image memory 272 to find a match. These images correspond to a series of images of the user in known positions, as illustrated in Figure 8D.

5 In Figure 8D, various images are stored representing various known positions of relevant part, here the user's head 256. In the position of Figure 8, for example a straight on face shot, the 0° stored memory image would provide the greatest cross-correlation value indicating a matched image position. Accordingly, the scene view would adjust to a zero position. If, however, the image correlated to a -90° position,
10 the scene would rotate to such a position. Other movements cause additional scene view motions, including tilt and tip of the head, as shown in the two images "0°, Down 45°" image and the "0°, Up 45°". These images cause the scene view to move upwards or to tilt up or down, when the process section 252 correlates the current frame to these images. As indicated, these images have no left or right component,
15 though other images (not shown) can certainly include left, right and tip motion simultaneously.

Figure 9 shows a system 300 constructed according to the invention and including a camera section 302 including an IR imager 304 and a camera 306, both of
20 which view and capture frames of data from a user 308. The IR imager 304 can include, for example, a microbolometer array (i.e., "uncooled" detectors known in the art) which produces a frame of data corresponding to the infrared energy emitted from the user, such as illustrated in Figure 9A. Figure 9A shows a representative frame of IR image data 310, with zones 312 of relatively hot image
25 data emitting from regions of forehead, nose and mouth of the user 308.

The cameras 304, 306 send image data back to the signal processing section 314. Data from the camera 306 is processed, if desired, as above, to determine difference information signal 322 used by a connected computer to reposition the
30 cursor and/or scene view. Data from camera 304, on the other hand, is used to evaluate how much (or how hot) zones 312 appear on the user during play of the

computer. The signal processing section 314 assesses the zones 312 for temperature and/or size over the course of a computer game and generate a "game speed control" signal 320 which is communicated to the user's computer (i.e., that computer used in conjunction with the system 300 of Figure 9). The user's computer
5 processes the signal 320 to increase or decrease the speed of a computer game in process on the computer.

Those skilled in the art should appreciate that the IR camera 304 can be used without the features of the invention which assess user movement. Rather, this
10 aspect should be considered stand-alone, if desired, to provide active feedback into gaming speed based upon user temperature and/or stress. Note that the camera 304 can also be used to detect heartbeat since the zones 312 generally pulse at the user's heartbeat, so that heartbeat rate can also be considered as a parameter used in the generation of the game speed control signal 320. Alternatively, a pulse rate can be
15 determined by known pulse rate systems that are physically connected to the user 308.

An IR lamp 324 can be used in system 300 to illuminate the user 308, with IR radiation 324a, such that sufficient IR illumination reflects off of the user 308
20 whereby motion control of the cursor and/or scene view can be made without the additional camera 306. The lamp 324 can be, and preferably is, made integrally with the section 302 to facilitate production packaging.

Figure 9B shows process methodology of the invention to process thermal
25 user images in accord with the preferred embodiment of the invention. Specifically, a system such as system 300 first acquires a thermal image map in process block 326. This image is compared to a reference image ("REF") in process block 327. REF can either be a temperature of the user (i.e., a temperature of one hot spot of a non-stressed user, or the temperature of one hot spot of the user at an initial, pre-game condition) or an amount of the area 312, Figure 9A, of the user in a non-stressed
30 condition or initial pre-game condition). By way of example, REF can be an image

such as the frame 310 of Figure 9A. When the temperature or area of the region 312 increases, the system 300 detects this change and determines that the image map exceeded the REF condition, as illustrated in process block 328. Should the map exceed the REF condition, the system 300 communicates this to the host processor
5 which in turn adjusts the gaming speed, as desired. If the map does not exceed the REF condition, then the next IR image frame is acquired at block 326.

System 300 and the process steps of Figure 9B are thus suitable to adjust gaming speed in real time, depending upon user stress level. In the preferred
10 embodiment, the gaming speed is increased automatically such that the image map exceeds the REF signal for greater than about 50% of the time, so that all users, regardless of their ability, are pushed in the particular game.

Those skilled in the art should appreciate that multi-camera embodiments of
15 the invention can and preferably are incorporated into a common housing 338, such as shown in Figure 10. Further, as illustrated in Figure 10, cameras can also be made from detector arrays 340, processing electronics 342, and optics 344. Each camera 340, 342, 344 is constructed to process the correct electromagnetic spectrum, e.g., IR (using, for example, germanium lenses 344 and microbolometer detectors 340). Each
20 camera has its own field of view 350a, 350b and focal distance 352a, 352b to image at least a part of the user. These field of views 350 can overlap, to view the same area such as the user's face, or they can view separate locations, such as the user's head and hand.

25 Cameras of the invention can also include a DSP section 356 such as described above to process user motion data. The DSP section 356 processes user motion data and sends difference information to the user's host computer. The host computer thereafter repositions the cursor and/or scene view based upon the difference information so that the user observes corresponding motion on the computer
30 display, as described above. Accordingly, the DSP section need not reside within the

computer so long as difference information is isolated and communicated to the host computer CPU.

Figure 11 illustrates frame capture by one camera of the invention to isolate zones of imaging according to expected motion patterns. In Figure 11, one frame 370 of data for example covers the user's eyes 371, corresponding to one image zone; and another frame 372 of data can cover the user's head 373, corresponding to another image zone. As mentioned previously, preferably the frames 370, 372 are 64x64 pixels each, or 256x256 (or higher 2 factorial set) to provide FFT capability on the image within the frame. A single camera can however provide both frames 370 and 372, in accord with the invention. Specifically, a dense CCD detector array (e.g., 480x740 pixels, 1000x1000 pixels, or higher) is used within the camera such that the whole array captures an image frame 376 of data, at least covering the available image format of the computer display 378. A matched filtering (or other image locate process) is processed on the frame 376 to locate the center 371a of the user's eyes (in the matched filtering process, an image data set of the user's eyes is stored in memory and correlated to the frame 376 such that a peak correlation is found at position 371a). Thereafter, a 64x64 array of data is centered about the eyes 371 to set the frame 370. To acquire the frame 372, every other pixel is discarded so that, again, a 64x64 array is set for the frame 372 (alternatively, each adjacent pair of pixels is added and averaged to provide a single number, again reducing the total number of pixels to 64x64). Note that this process is reasonable since the width of the eyes is at least $\frac{1}{2}$ the width of the user's face. Nevertheless, further compression can be obtained by utilizing every third pixel (or averaging three adjacent pixels) to obtain a larger image area in the frame 372. Note that the compression in the width and length dimensions need not be the same.

Framing of the information in Figure 11 can occur in several ways. Most cameras image at 30Hz so that image motion is smooth to the human eye. In one embodiment, one frame 370 is taken in between each frame 372, to minimize data throughput and processing; and yet to maintain dual processing of the two zones

imaged in Figure 11. Alternatively, both frames 370, 372 are processed concurrently since frame 376 is typically the 30Hz frame.

Figure 11 also illustrates how framing can occur around the user's eyes 371 to acquire "blink" information to reset cursor control. A blink detected by the user's eyes in frame 370 (or other frame) can be used to (a) disable or enable control of cursor or scene movement based upon user control, or (b) simulate pick-up and replacement of the computer mouse (i.e., reinitializing movement in a particular direction). For example, by detecting a blink of the eyes 371, a system of the invention can disable human motion following control such as described herein. Another blink can be used to enable human motion following control. Blinking can also be used to continue motion in a particular direction. For example, movement of the cursor can be made to follow movement of the user's head, as described above. However, after a while, the person has to move to an uncomfortable position to keep moving the cursor (or scene). A blink can thus also serve to reposition the head back to a normal starting position so that further movement in the desired direction can be made.

Figure 12 illustrates a similar capture of a user's eyes 400, in accord with the invention. A frame 402 can thus be acquired by a camera of the invention. Figure 12A illustrates further detail of one representative frame 402, illustrating that the user's pupils 404 are also captured. Figures 3 and 4 describe certain algorithms of the invention that are also applicable to motion of the user's pupils 404, as illustrated by left and right motion 406 and up and down motion 408. Accordingly, by zooming in on the user's eyes, another movement zone is created that causes repositioning of the cursor or the scene view based upon the movements 406, 408, much like the head movement described and illustrated in Figures 1-4.

Note that the teachings of Figures 1-4 and 12-12A can be combined within a two zone movement system so that, for example, both head motion and pupil motion can be evaluated for image motion. The cursor and/or scene view can be

repositioned, therefore, based upon movements from both zones. By way of example, repositioning of items within the display (e.g., the cursor and/or scene view) can be made when the head moves but not if the head and eyes move, which might indicate that the user is simply looking elsewhere in the room due to a distraction. However, if the user moves his head, but not his eyes, he is focussed on the game and intends rotation of the scene view, in another example. Other combinations are also possible.

Cameras of the invention can also include zoom optics which (a) reduce or enlarge the image frame captured by a particular camera, or which (b) provide autofocus capability. Figure 13 shows one system 430 constructed according to the invention. A camera 432 includes camera electronics 432a and a zoom attachment 432b. Data from the camera 432 is relayed to image and interpretation feedback electronics 434 for evaluation. For purposes of image magnification, the feedback electronics serve to evaluate a given image size relative to desired image goals. For example, to image the user's eyes with high fidelity might require high density of pixels at the user's eyes (e.g., at the zone 370, Figure 11). Accordingly, the system 430 can isolate the user's eyes, such as described herein, and command the camera (through command lines 436) to increase or decrease magnification on the user's eyes so as to achieve desired resolution. The feedback electronics can also command motion of the camera to change its boresight alignment (i.e., to change where the camera image is centered) by commanding movement of the camera when resting on one or more linear drives 438, as known in the art.

Once aligned on the desired user location, e.g., on the eyes with desired accuracy, the system 430 continues processing data such as described herein to create human interface control of items displayed on the user's computer, e.g., cursor and/or scene view. Accordingly, processing section 440 operates to detect user motion and to communicate difference information to the user's computer, as described above.

With regard to autofocus, the system 430 of Figure 13 can also be used to process user motion based upon motion towards and away from the camera. Figure 14 illustrates such a system, including a camera 450 with autofocus capability to find the best focus 452 relative to a user 454 within the field of view 456. For example, when the user 454 moves to position 460 (the user being shown in outline form 454a), the new best focus has changed to 452a. The camera 450 provides a signal 450a to the image interpretation and feedback electronics 434, Figure 13, which indicates where the user is along the "z" axis from the camera 450 to the user 454. This signal 450a is thus used much like the other motion signals described herein, to move the cursor and/or scene view in response to such movements. Figure 14A illustrates a representative scene view 462 when, for example, the user is at best focus 452. The scene view 462 includes a house image 464 with a door 465. When the user moves to position 460, the house and door 464', 465' of the scene view 462' enlarge, since the user moved closer to the camera 450. Such a motion might reveal, for example, additional objects within the house, such as illustrated by object 466, Figure 14B. Accordingly, the autofocus feature of the invention provides yet another degree of freedom in motion control, in accord with the invention.

Image data, manipulation, and human interface control can be improved, over time, by using neural net algorithms. As shown in Figure 13, a neural net update section 435 can for example couple to the feedback electronics 434 so as to assimilate movement information and to improve data transmitted to the host computer, over time. Use of neural nets are known in the art.

Figure 15 illustrates a frame of data 490 used in accord with the invention to implement a simplified left, right, up, down movement algorithm to control cursor movement and/or scene view movement. Frame 490 is captured by a camera of the invention; and preferably the camera incorporates autofocus, as described above, to provide a crisp image of the user 492 regardless of her position within the camera's field of view. As shown, image frame 490 provides very sharp edges to the user's face, including a left edge 494a, right edge 494b, and chin 494c. These edges need

only approximate vertical or horizontal position. Movement of the user results in movement of the edges 494, such as shown in Figure 15A. Figure 15A shows that once such edges are acquired, they conveniently permit subsequent movement analysis and control of scene view and/or cursor position. Specifically, Figure 15A shows movement of the user's "edges" from 494a-c to 494a-c', indicating that the user moved left (as viewed from the camera's position) and that her chin raised slightly, indicating that an upward tilt of the head. This information is assessed by the process sections as discussed above and relayed to the host computer as difference information to augment or provide cursor and/or scene movement in response to the user's movement.

Note that such edge movements roughly correspond to movement along rows and columns of the detector array. Detected movement from one row to another (or one column to another) can readily calculate the actual motion of the user from information of the user's best focus position and from the focal length of the camera's lens. This information may then be used to set the magnification of movement of items in the computer display (e.g., cursor and/or scene view).

Figure 16 illustrates an image of one object 500 used in accord with the invention to provide image manipulation in response to motion of the object. The object 500 is held by the user 501 to manipulate motion of his cursor 502 and/or scene view 504 on his computer display 506. The object 500 is used because it exhibits an optical shape that is easily recognized through image correlation (such as matched filtering). In accord with the invention, a camera 510 is used to image the object 500; and frames of data are sent to the frame processor 512. The processor 512 determines image position - relative to a starting position - and thereafter communicates difference information to the user's computer 505 along data line 514. The difference information is used by the computer's CPU and operating system to reposition items on the display 506 in response to motion of the object 500. Almost any motion, including rotation, tilting and translation are accomplished with the object 500 relative to a start position. This start position can be triggered by the user

501 at the start of a game by commanding that the camera 510 take a reference frame ("REF") that is stored in memory 513. The user 501 commands that REF imagery be taken and stored through the keyboard 505a, connected to the computer 505, which in turn commands the processor 512 and camera 510 to take the reference frame REF.

5

Motion of the object 500 is thus made possible with enhanced accuracy by comparing subsequent frames of the object 500 with REF. When motion of rotation, tilt or translation are detected (for example, by using the techniques of Figures 2-4, 8-8D), then repositioning of items (502, 504) on the display 506 are follow that movement.

10

The techniques of the invention permit control of the scene view and/or cursor on a computer screen by motion of one or more parts of the user's body. Accordingly, as shown in Figure 17, complete motion of the user 598 can be replicated, in the invention, by correlated motion of a action figure 599 within a game. In Figure 17, user 598 is imaged by a camera 602 of the invention; and frames from the camera 602 are processed by the process section 604, such as described herein. The user 598 is captured and processed, in digital imagery, and annotated with appropriate user segments, e.g., segments 1-6 indicating the user's hands, feet, head and main body. Motion of the segments 1-6 are communicated to the host computer 606 from the process section 604. The computer's operating system then updates the associated display 608 so that the action figure 599 (corresponding to an action figure within a computer game) moves like user 598. Accordingly, user motion of action figure 599 is made by the user 598 by performing stunts (e.g., striking and kicking) that he would like the action figure 599 to perform, such as to knock out an opponent within the display 608.

20

25

30

Figure 18 illustrates a two camera system 700 used to determine translation and rotation. The forward viewing camera 702 observes the user's face 703 and determines the right-left (Δx_1) and up-down (Δy_1) translation of the user's face 793. The top viewing camera 704 observes the top of the user's face or head 705 and

determines the right-left (Δx_2) and up-down motion (Δy_2) of the user's face or head. The two cameras 702, 704 are each processed through motion sensing algorithms 706 using the teachings above, and results are shown on the computer display 710. For purposes of illustration, the display 710 shows an image of the user; while the
5 image can be, for example, an action figure or other computer object, as desired. As indicated in Figure 18, the Δy_2 can be directly applied to motion control of the user's forward and reverse motion (note, these motions are illustrated as within a computer display 710 as processed by algorithms 706). Δx_2 can be directly applied to the users left-right sideways or strafe motion; Δy_1 can be directly applied to control
10 the users up-down viewpoint, each as illustrated on display 710a. The results of the difference between Δx_2 and Δx_1 can be applied to control the user's left-right turn or viewpoint.

The techniques of Figure 18 can be further extended to front, side and top
15 view cameras for complete motion detection. The top camera determines the user's left-right, front-back motion while the front facing camera determines the user's rotational up-down, left-right motion.

Figure 19 describes an algorithm to detect user eye blink. The video imagery
20 is stored into a multiple frame buffer 800. The algorithm selects the current frame and a frame from the frame buffer and differences these frames using the adder 802. The difference frame consists of the pixel by pixel difference of the delayed frame and the current frame. The difference frame includes motion information used by the algorithms of teachings above. It also contains information on the user eye blink.
25 The frames differenced by the adder 802 are separated temporally enough to ensure that one frame contains an image of the users face with the eyes open, the other image is of the user's face with the eyes closed. The difference image contains a two strong features, one for each eye. These features are spatially separated by the distance between the user's eyes. The blink detect function 808 inspects the image
30 for this pair of strong features which are aligned horizontally and spaced within an expected distance based on the variation from one human face to another and the

variation in seating distance expected from user to user. The recognition of the blink features may be accomplished using a matched filter or by recognition of expected frequency peaks in the frequency domain at the expected spatial frequency for human eye separation. The blink detect function 708 identifies the occurrence of a
5 blink to a controlling function to either disable the cursor motion or take some other action.

Appendix A, pages A1-A27, provide non-limiting source code to illustrate certain features and operations of the invention.

10

The invention thus attains the objects set forth above, among those apparent from the preceding description. Since certain changes may be made in the above methods and systems without departing from the scope of the invention, it is intended that all matter contained in the above description or shown in the
15 accompanying drawing be interpreted as illustrative and not in a limiting sense. It is also to be understood that the following claims are to cover all generic and specific features of the invention described herein, and all statements of the scope of the invention which, as a matter of language, might be said to fall there between.

20

Having described the invention, what is claimed is:

1. A human motion following controller for augmenting motion of items shown on a computer display, the display being coupled to a computer of the type which controls positioning of the items through operating system controls, comprising:

5 a camera for capturing frames of data corresponding to an image of a first part of a user at the computer display;

signal processing means coupled to the camera for (a) detecting differences between successive frames of data corresponding to motion of the first part, and (b)
10 communicating differences information to the computer to reposition display of the items through operating system controls, the items being repositioned on the display by an amount corresponding to the motion of the first part.

2. A controller of claim 1, wherein the items comprise a computer cursor.

15

3. A controller of claim 1, wherein the items comprise a scene view.

4. A controller of claim 1, further comprising a PC card for installation within the computer and for communication on a computer bus, the signal processing
20 means being substantially resident with the PC card for communicating differences information to the bus.

5. A controller of claim 1, wherein the camera comprises means for capturing augmented frames of data corresponding to a second part of a user at the computer
25 display, the signal processing means further comprising means for detecting differences between successive augmented frames of data corresponding to motion of the second part and for communicating augmented differences information to the computer to reposition display of the items through operating system controls, the items being repositioned on the display by an amount corresponding to motion of
30 the first and second parts.

FIGURE 1

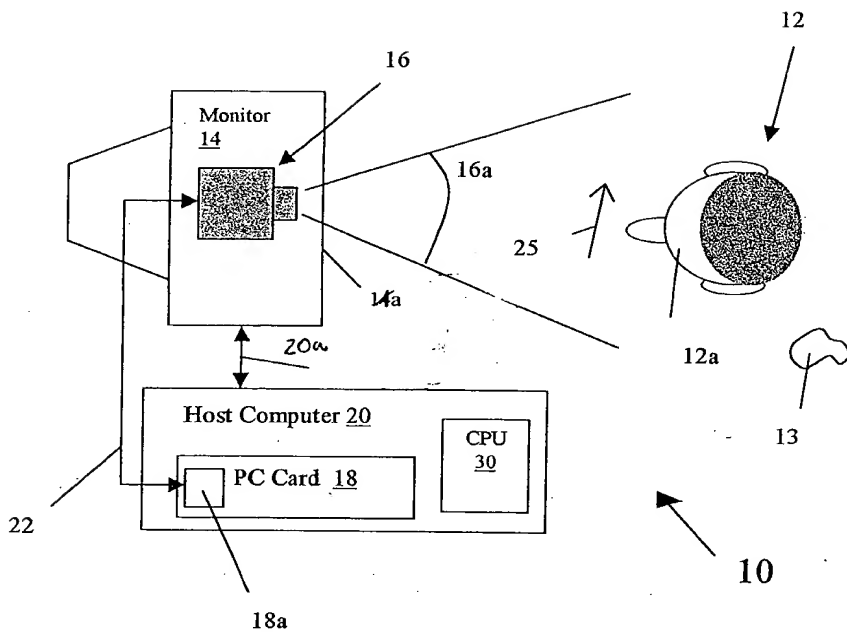


FIGURE 1A

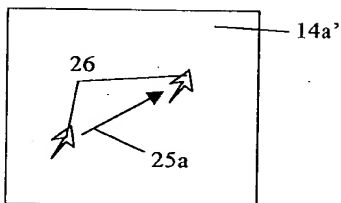


FIGURE 1B

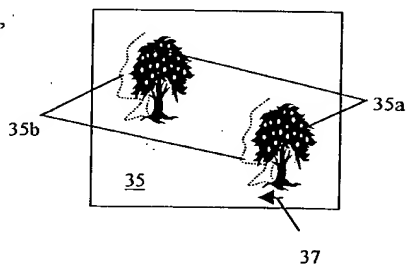


FIGURE 1C

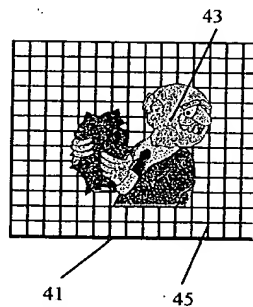


FIGURE 2

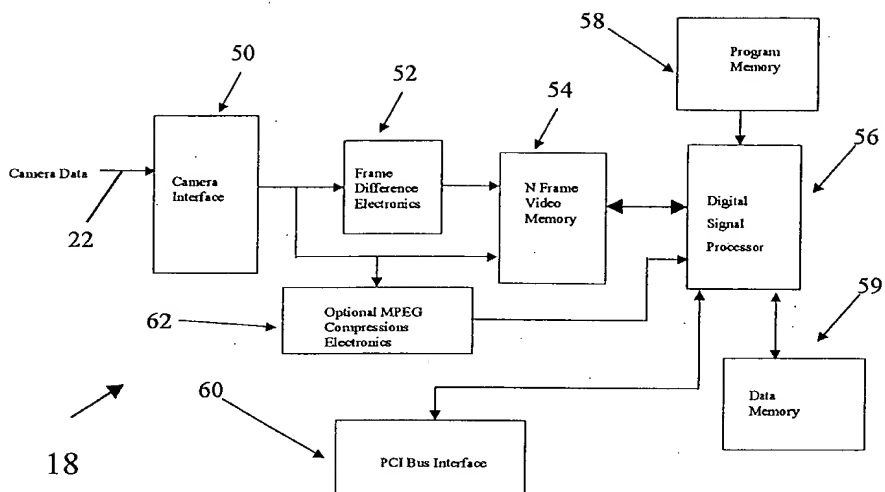


FIGURE 3

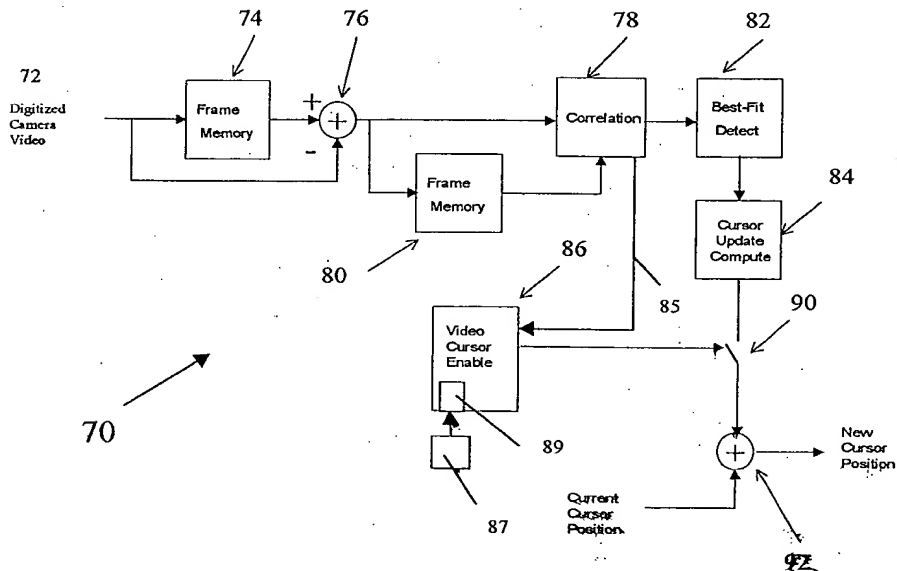


FIGURE 4

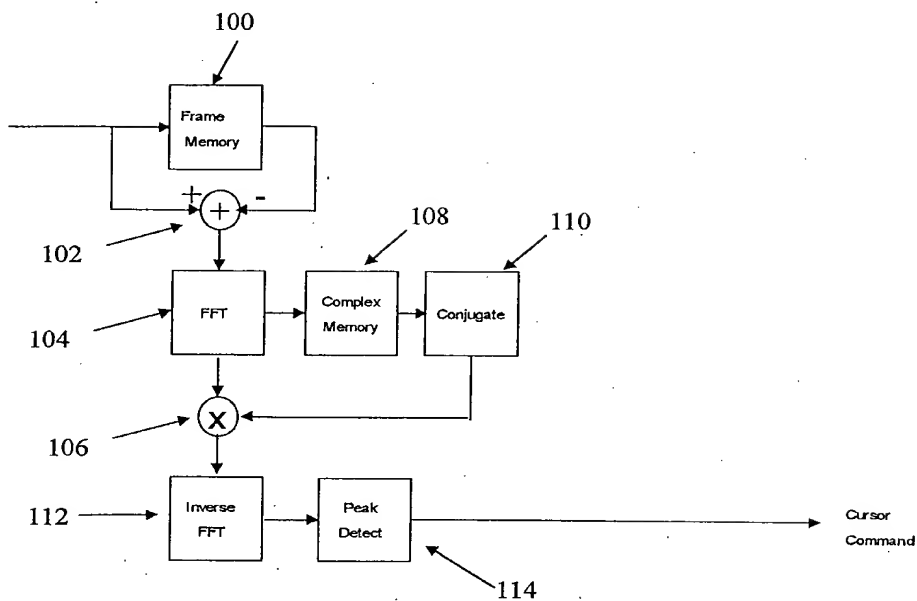


FIG. 5

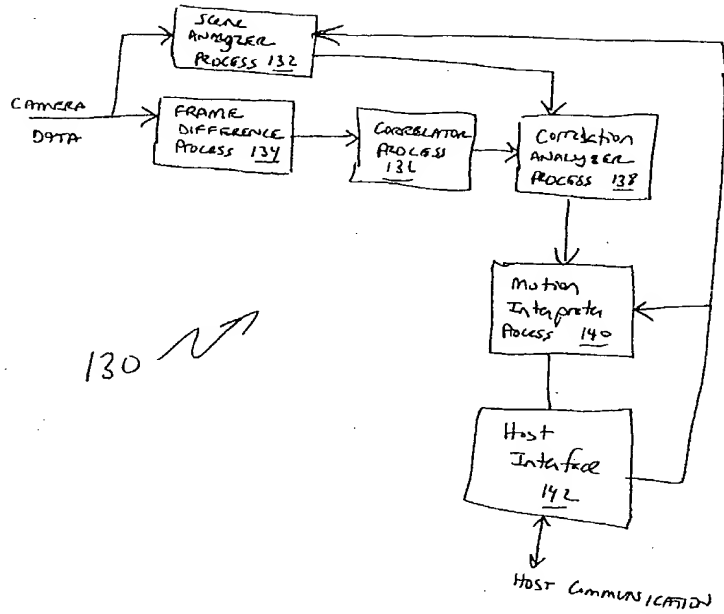
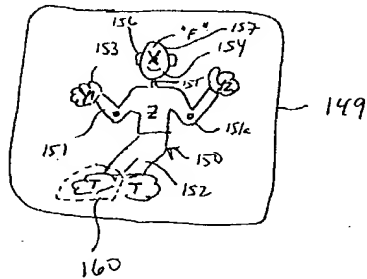
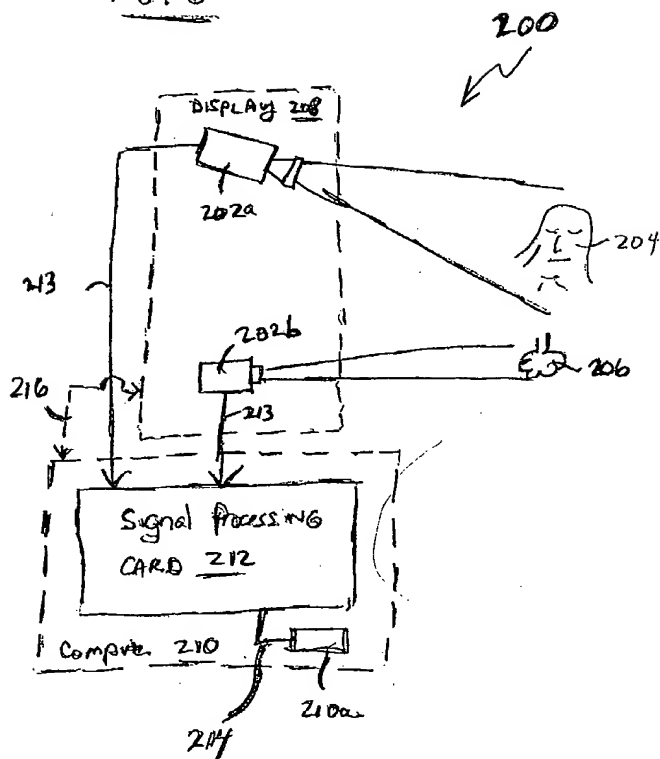


FIG. 5A



4/17

FIG. 6



5/17

FIG. 7

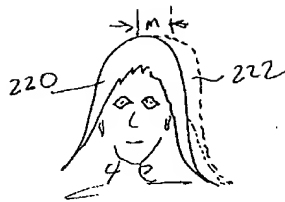


FIG. 7A

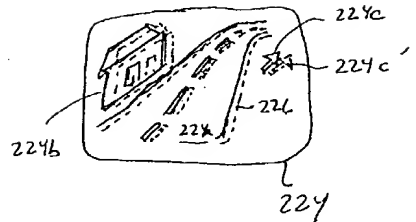


FIG. 8

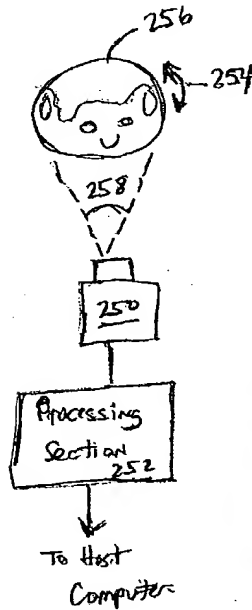


FIG. 8A

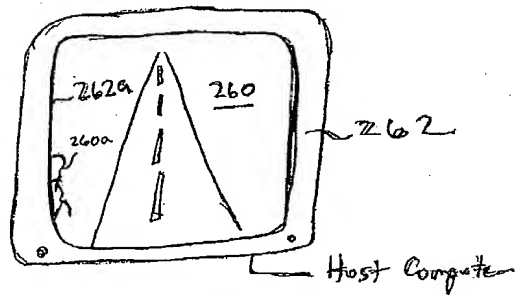


FIG. 8B

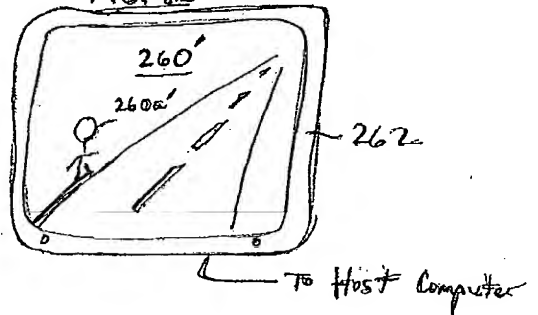
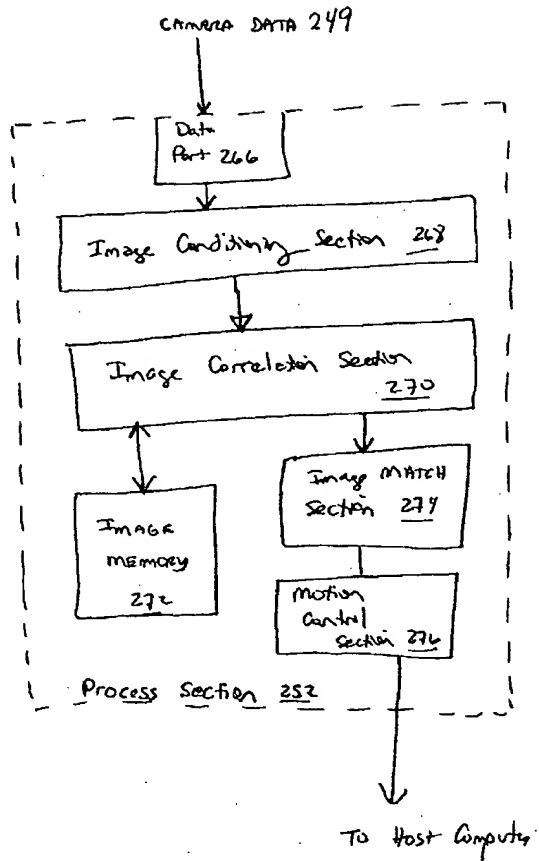
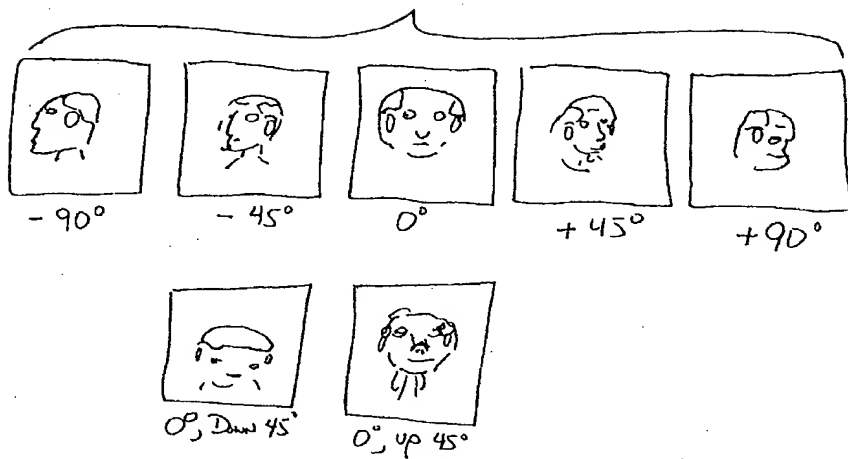


FIG. 8C



7/17

FIGURE 8D



8/17

FIG. 9

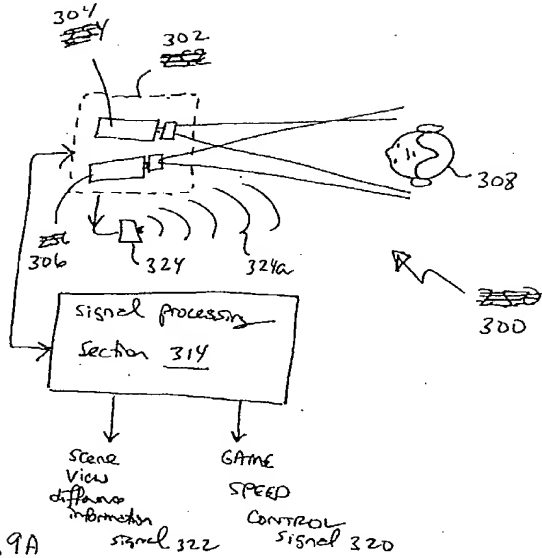
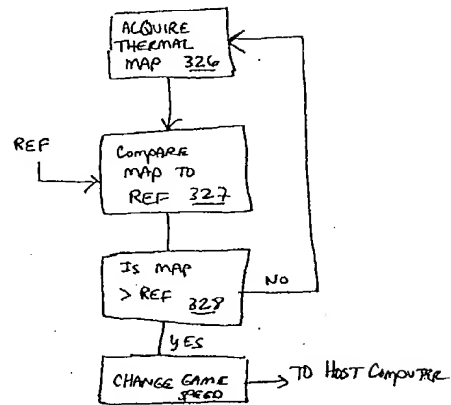


FIG. 9A



FIG. 9B



9/17

FIG. 10

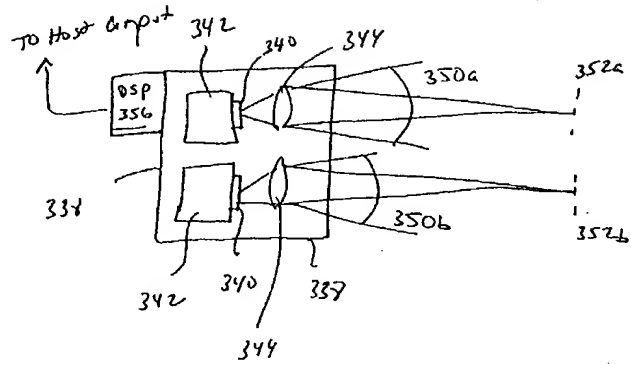
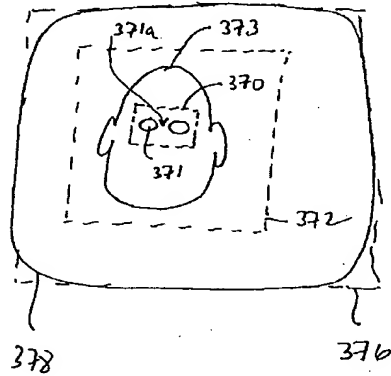


FIG. 11



10/17

FIG. 12

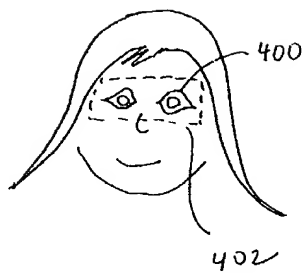


FIG. 12A

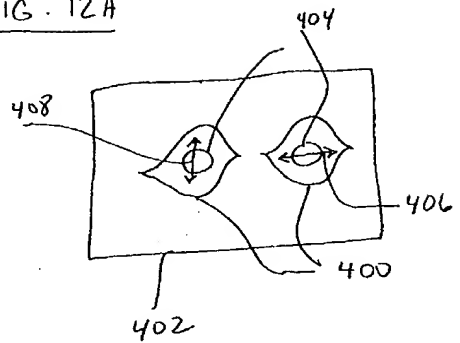


FIG. 13

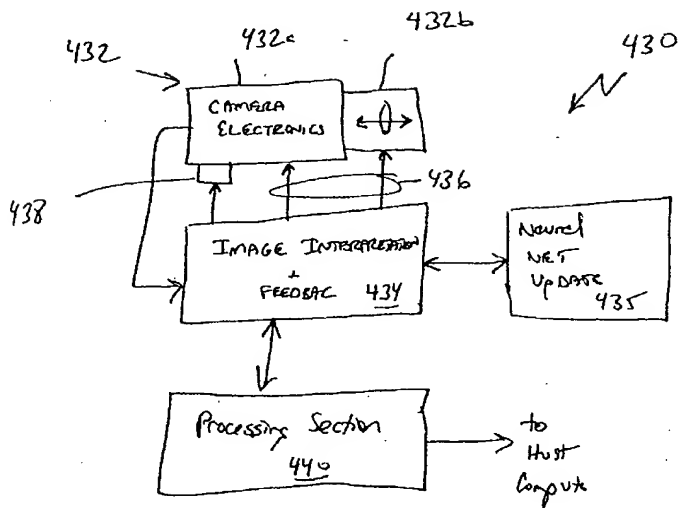


FIG. 14

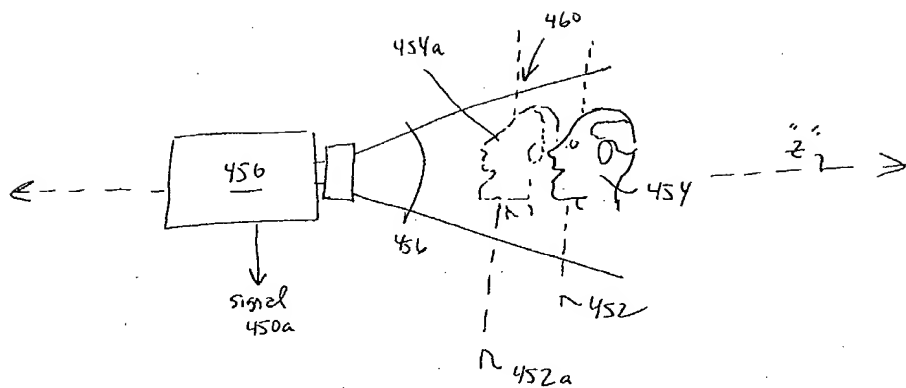


FIG. 14A

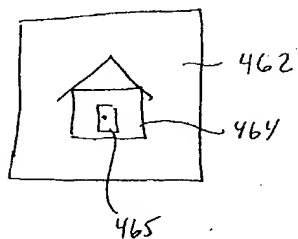


FIG. 14B

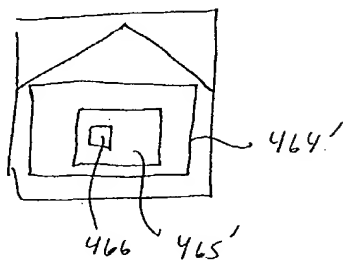


FIG. 15

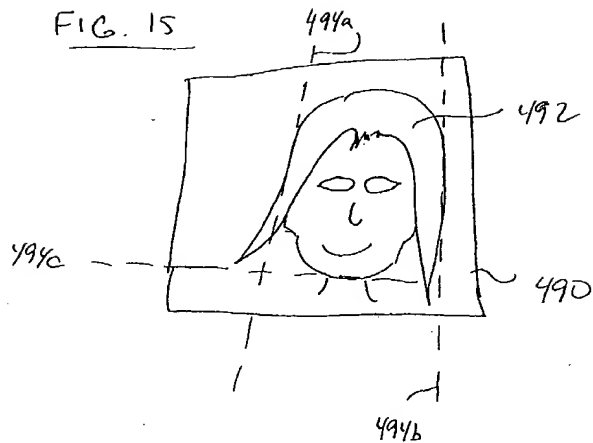


FIG. 15A

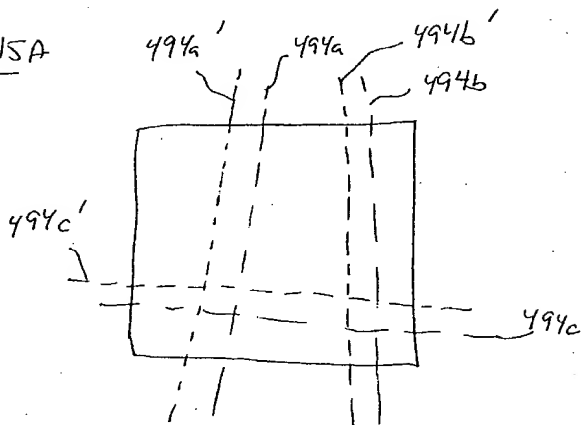
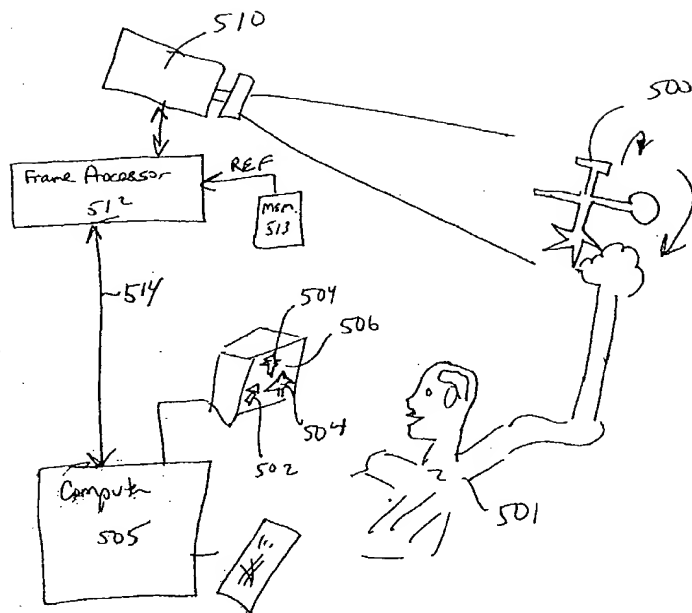
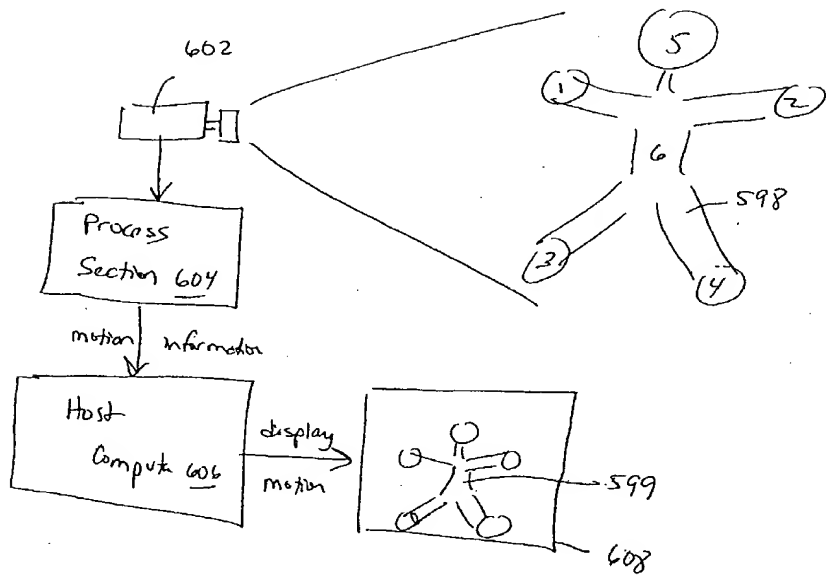


FIG. 16



14/17

FIG. 17



15/17

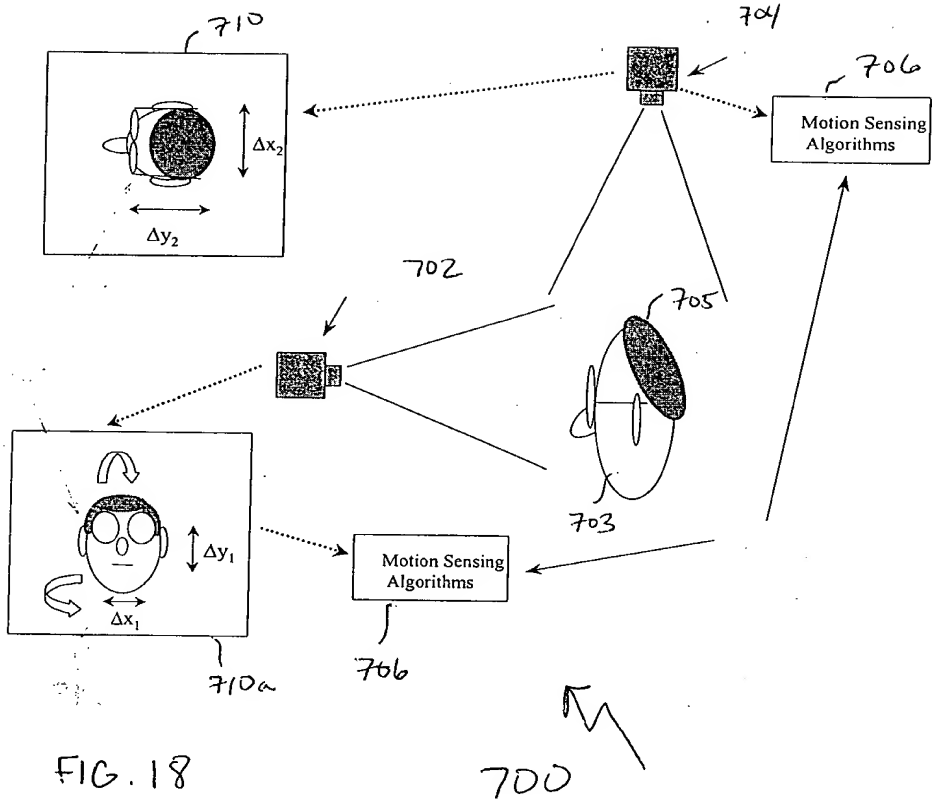
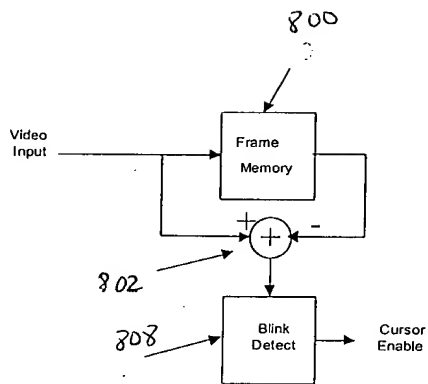


FIG. 18

FIG. 19




```
// cursor.cpp Demonstration program for video mouse.
#include "glos.h"

#include <stdio.h>
5  #include <string.h>
#include <stdlib.h>
#include <math.h>
#include <gl\glaux.h>
#include "fftn.h"

10  #include "mv1.h"

#define NORMAL 0
#define FRAMESIZE 64
15  #define ZOOM 2.0
#define DISPLAY 1

static void CALLBACK Key_q(void);
static void CALLBACK CursorControl(void);
20  void MTPROCALL caminterrupt(short);
void InitializeCamera(void);
void StartFrameGrab(void);
void PositionCursor(double detectx, double detecty);

25  int numPixels;
unsigned char *camBuffer;
double *reData1, *imData1;
double *reData2, *imData2;
double *reCorrelation, *imCorrelation;
30  unsigned char *camFrame;

volatile long waitforinterrupt;

35  void main(int argc, char **argv)
{
    int windW = ZOOM*FRAMESIZE;
    int windH = ZOOM*FRAMESIZE;
40  auxInitPosition(0, 0, windW, windH);

    auxInitDisplayMode(AUX_RGB|AUX_DOUBLE);

    if (auxInitWindow("Headmaus View") == GL_FALSE) {
45  auxQuit();
    }

    glClearColor(0.0, 0.0, 0.0, 0.0);
    glPixelZoom(ZOOM, ZOOM);
50  InitializeCamera();

    auxKeyFunc(AUX_q, Key_q);
55  // auxIdleFunc(NULL);
    auxMainLoop(CursorControl);
}
```

```
static void CALLBACK CursorControl(void)
{
    int k;
    int even=1;
    unsigned char temp = 0;
    int dims[2]={FRAMESIZE,FRAMESIZE};
    double multiplier;

    while(1){
    while(waitforinterrupt==0);
    waitforinterrupt=1;
    if (even){
        even=0;
        for (k=0;k<numPixels;k++){
            temp=camBuffer[numPixels-k-1];
            reData1[k] = (double) (camFrame[k]-temp);
            imData1[k] = 0.0;
            camFrame[k]=temp;
        }
        fftn(2,dims,reData1, imData1,1,0);
        for (k=0;k<numPixels;k++){
            reCorrelation[k] = reData1[k]*reData2[k]+imData1[k]*imData2[k];
            imCorrelation[k] = imData1[k]*reData2[k]-reData1[k]*imData2[k];
        }
    }
    else {
        even=1;
        for (k=0;k<numPixels;k++){
            temp=camBuffer[numPixels-k-1];
            reData2[k] = (double) (camFrame[k]-temp);
            imData2[k] = 0.0;
            camFrame[k]=temp;
        }
        fftn(2,dims,reData2, imData2,1,0);
        for (k=0;k<numPixels;k++){
            reCorrelation[k] = reData2[k]*reData1[k]+imData2[k]*imData1[k];
            imCorrelation[k] = imData2[k]*reData1[k]-reData2[k]*imData1[k];
        }
    }

    fftn(2,dims,reCorrelation, imCorrelation,-1,-1);

    double max=-200000.0;
    int maxindex=0;

    for (k=0;k<numPixels;k++){
        if (reCorrelation[k] > max){
            max=reCorrelation[k];
            maxindex=k;
        }
    }

    glClear(GL_COLOR_BUFFER_BIT);
    glDrawPixels(FRAMESIZE, FRAMESIZE, GL_LUMINANCE,
    GL_UNSIGNED_BYTE,camFrame);
    glFlush();
    auxSwapBuffers();
}
```

```
        if (max>5000.0){
            double detecty=(double)((int)(maxindex/FRAMESIZE));
            double detectx=(double)(maxindex - detecty*FRAMESIZE);

5            if (detectx > FRAMESIZE/2) detectx = detectx - FRAMESIZE;
            if (detecty > FRAMESIZE/2) detecty = FRAMESIZE - detecty;
            else detecty = -detecty;

            double vector=sqrt(detectx*detectx+detecty*detecty);
10            if (vector<2) multiplier=1;
            else if (vector<10) multiplier = exp((vector-2)/2.5);
            else multiplier=0;

            detectx*=multiplier;
15            detecty*=multiplier;

            PositionCursor(detectx,detecty);
        }
    }
20 }
static void CALLBACK Key_q(void)
{
    MVIStopGrab();
25    MVIIRQDisable(MV1_MXfer_Int);
    MVIDisconnectInterruptCallback(0);

    MVIUnlockMasterBuffer(camBuffer, numPixels);

30    fft_free();
    free(imData1);
    free(reData1);
    free(imData2);
    free(reData2);
35    free(imCorrelation);
    free(reCorrelation);
    free(camBuffer);
    free(camFrame);
    auxQuit();
40 }
void MTPROCALL caminterrupt(short status)
{
    MVIInterruptProcessed(0);
45    waitforinterrupt = 0;
}
void InitializeCamera(void)
{
50    MVIGrabWindow grabWindow;
    MVIFrame Frame;
    unsigned long camrows, camcols;
    int mvstatus=MVIOpen();
    mvstatus=MV1SetCurrentBoard(0);
55    mvstatus=MV1LoadCameraConfigFile("f:\\mv-1000\\camcfg\\headmouse.ini", "CamConfig");
    MVIInquireMaxGrabWindowSize(&camcols,&camrows);
    numPixels = FRAMESIZE*FRAMESIZE;
    camBuffer = (unsigned char *)malloc(numPixels*sizeof(unsigned char));
```

```
    camFrame = (unsigned char *) malloc(numPixels*sizeof(char));
    //FFT Stuff
    reData1 = (double *) malloc(numPixels*sizeof(double));
    imData1 = (double *) malloc(numPixels*sizeof(double));
5
    reData2 = (double *) malloc(numPixels*sizeof(double));
    imData2 = (double *) malloc(numPixels*sizeof(double));
    reCorrelation = (double *) malloc(numPixels*sizeof(double));
    imCorrelation = (double *) malloc(numPixels*sizeof(double));
10
    // end FFT Stuff

    int rowstart=camrows/2-FRAMESIZE/2;
    int colstart=camcols/2-FRAMESIZE/2;

15
    long windowaddress=0x10000;
    MV1SetGrabWindow(windowaddress, colstart, rowstart, FRAMESIZE, FRAMESIZE,
    &grabWindow) ;

    MV1CreateFrame(&Frame, 0, 0, FRAMESIZE, FRAMESIZE, &grabWindow) ;
20
    MV1ConnectInterruptToCallback(0, (FARPROC) &caminterrupt, MV1_MXfer_Int);

    MV1SelectTransferMode(MV1_Master_Mode | MV1_Master_Loop_Mode |
    MV1_Master_FrameEnd_Mode);
25
    MV1LockMasterBuffer(camBuffer, numPixels);

    MV1SetMasterModeTransfer(&Frame, camBuffer, numPixels, 0);
    MV1SetMasterCtlId(0);
30
    MV1SelectTransferMode(MV1_Master_Mode | MV1_Master_Loop_Mode |
    MV1_Master_FrameEnd_Mode);

    MVIIRQEnable(MV1_MXfer_Int);
35
    MV1StartGrab(MV1_Cont_Grab, MV1_Grab_Even) ;
    waitforinterrupt = 1;
}
void PositionCursor(double detectx, double detecty)
{
40
    static POINT ptCursor;
    GetCursorPos(&ptCursor);

    ptCursor.x+=(long) detectx;
    ptCursor.y+=(long) detecty;
45
    SetCursorPos(ptCursor.x,ptCursor.y);
}
```

```

/*-----*-C-*-----*
* File:
*     fftn.c
*
5  *     multivariate complex Fourier transform, computed in place
*     using mixed-radix Fast Fourier Transform algorithm.
*
*     Fortran code by:
*         RC Singleton, Stanford Research Institute, Sept. 1968
10  *         NIST Guide to Available Math Software.
*         Source for module FFT from package GO.
*         Retrieved from NETLIB on Wed Jul 5 11:50:07 1995.
*         translated by f2c (version 19950721) and with lots of cleanup
*         to make it resemble C by:
15  *         MJ Olesen, Queen's University at Kingston, 1995-97
*     */
/*{{{ Copyright: */
/*
20  * Copyright(c)1995,97 Mark Olesen <olesen@mc.QueensU.CA>
*         Queen's Univ at Kingston (Canada)
*
* Permission to use, copy, modify, and distribute this software for
* any purpose without fee is hereby granted, provided that this
* entire notice is included in all copies of any software which is
25  * or includes a copy or modification of this software and in all
* copies of the supporting documentation for such software.
*
* THIS SOFTWARE IS BEING PROVIDED "AS IS", WITHOUT ANY EXPRESS OR
* IMPLIED WARRANTY. IN PARTICULAR, NEITHER THE AUTHOR NOR QUEEN'S
30  * UNIVERSITY AT KINGSTON MAKES ANY REPRESENTATION OR WARRANTY OF ANY
* KIND CONCERNING THE MERCHANTABILITY OF THIS SOFTWARE OR ITS
* FITNESS FOR ANY PARTICULAR PURPOSE.
*
* All of which is to say that you can do what you like with this
35  * source code provided you don't try to sell it as your own and you
* include an unaltered copy of this message (including the
* copyright).
*
* It is also implicitly understood that bug fixes and improvements
40  * should make their way back to the general Internet community so
* that everyone benefits.
*-----*/
/*}}}*/
/*{{{ notes: */
45  /*
* Public:
*     fft_free
*     fftn / fftmf
*     (these are documented in the header file)
50  *
* Private:
*     fftadix / fftadixf
*-----*
55  * int fftadix (REAL Re[], REAL Im[], size_t nTotal, size_t nPass,
*             size_t nSpan, int iSign, size_t maxFactors,
*             size_t maxPerm);
*
* RE and IM hold the real and imaginary components of the data, and
    
```

```

* return the resulting real and imaginary Fourier coefficients.
* Multidimensional data *must* be allocated contiguously. There is
* no limit on the number of dimensions.
*
5  *
* Although there is no limit on the number of dimensions, fftadix()
* must be called once for each dimension, but the calls may be in
* any order.
*
10 * NTOTAL = the total number of complex data values
* NPASS = the dimension of the current variable
* NSPAN/NPASS = the spacing of consecutive data values while indexing
* the current variable
* ISIGN - see above documentation
15 *
* example:
* tri-variate transform with Re[n1][n2][n3], Im[n1][n2][n3]
*
*      fftadix (Re, Im, n1*n2*n3, n1, n1, 1, maxf, maxp);
20 *      fftadix (Re, Im, n1*n2*n3, n2, n1*n2, 1, maxf, maxp);
*      fftadix (Re, Im, n1*n2*n3, n3, n1*n2*n3, 1, maxf, maxp);
*
* single-variate transform,
* NTOTAL = N = NSPAN = (number of complex data values),
25 *
*      fftadix (Re, Im, n, n, n, 1, maxf, maxp);
*
* The data can also be stored in a single array with alternating
* real and imaginary parts, the magnitude of ISIGN is changed to 2
30 * to give correct indexing increment, and data [0] and data [1] used
* to pass the initial addresses for the sequences of real and
* imaginary values,
*
* example:
35 *      REAL data [2*NTOTAL];
*      fftadix (&data[0], &data[1], NTOTAL, nPass, nSpan, 2, maxf, maxp);
*
* for temporary allocation:
*
40 * MAXFACTORS      >= the maximum prime factor of NPASS
* MAXPERM      >= the number of prime factors of NPASS. In addition,
* if the square-free portion K of NPASS has two or more prime
* factors, then MAXPERM >= (K-1)
*
45 * storage in FACTOR for a maximum of 15 prime factors of NPASS. if
* NPASS has more than one square-free factor, the product of the
* square-free factors must be <= 210 array storage for maximum prime
* factor of 23 the following two constants should agree with the
* array dimensions.
50 * -----*/
/*}}}}*/
/*{{{ Revisions: */
/*
* 26 July 95      John Beale
55 *      - added maxf and maxp as parameters to fftadix()
*
* 28 July 95      Mark Olesen <olesen@mc.QueensU.CA>
*      - cleaned-up the Fortran 66 goto spaghetti, only 3 labels remain.
*

```

```

*      - added fft_free() to provide some measure of control over
*      allocation/deallocation.
*
*      - added fftm() wrapper for multidimensional FFTs
5  *
*      - use -DFFT_NOFLOAT or -DFFT_NODOUBLE to avoid compiling that
*      precision. Note suffix 'f' on the function names indicates
*      float precision.
*
10 *      - revised documentation
*
* 31 July 95      Mark Olesen <olesen@me.QueensU.CA>
*      - added GNU Public License
*      - more cleanup
15 *      - define SUN_BROKEN_REALLOC to use malloc() instead of realloc()
*      on the first pass through, apparently needed for old libc
*      - removed #error directive in favour of some code that simply
*      won't compile (generate an error that way)
*
20 * 1 Aug 95      Mark Olesen <olesen@me.QueensU.CA>
*      - define FFT_RADIX4 to only have radix 2, radix 4 transforms
*      - made ffradix /fftradixf () static scope, just use fftm()
*      instead. If you have good ideas about fixing the factors
*      in fftm() please do so.
25 *
* 8 Jan 95      mj olesen <olesen@me.QueensU.CA>
*      - fixed typo's, including one that broke scaling for scaling by
*      total number of matrix elements or the square root of same
*      - removed unnecessary casts from allocations
30 *
* 10 Dec 96      mj olesen <olesen@me.QueensU.CA>
*      - changes defines to compile *without* float support by default,
*      use -DFFT_FLOAT to enable.
*      - shifted some variables to local scope      (better hints for optimizer)
35 *      - added Michael Steffens <Michael.Steffens@mbx.muk.uni-hannover.de>
*      Fortran 90 module
*      - made it simpler to pass dimensions for 1D FFT.
*
* 23 Feb 97      Mark Olesen <olesen@me.QueensU.CA>
40 *      - removed the GNU Public License (see 21 July 1995 entry),
*      which should make it clear why I have the right to do so.
*      - Added copyright notice and submitted to netlib
*      - Moved documentation for the public functions to the header
*      files where is will always be available.
45 * -----*/
/*}}}*/
#ifndef FFTN_C
#define FFTN_C
/* we use CPP to re-include this same file for double/float cases */
50 #if !defined (lint) && !defined (__FILE__)
Error: your compiler is sick! define __FILE__ yourself (a string)
eg, something like -D__FILE__="fft.c"
#endif
#include <stdio.h>
55 #include <stdlib.h>
#include <math.h>
#include "fftn.h"

/*{{{ defines/constants */
```

```

#ifndef M_PI
#define M_PI 3.14159265358979323846264338327950288
#endif

5  #ifndef SIN60
    #define SIN60 0.86602540378443865 /* sin(60 deg) */
    #define COS72 0.30901699437494742 /* cos(72 deg) */
    #define SIN72 0.95105651629515357 /* sin(72 deg) */
    #endif
10 /*}}}*/

    /*{{{ static parameters - for memory management */
    static size_t SpaceAlloced = 0;
    static size_t MaxPermAlloced = 0;
15 /* temp space, (void *) since both float and double routines use it */
    static void * Tmp0 = NULL; /* temp space for real part */
    static void * Tmp1 = NULL; /* temp space for imaginary part */
    static void * Tmp2 = NULL; /* temp space for Cosine values */
20 static void * Tmp3 = NULL; /* temp space for Sine values */
    static int * Perm = NULL; /* Permutation vector */

#define NFACTOR 11
    static int factor[NFACTOR];
25 /*}}}*/

    /*{{{ fft_free() */
    void
    fft_free(void)
30 {
        SpaceAlloced = MaxPermAlloced = 0;
        if (Tmp0) { free (Tmp0); Tmp0 = NULL; }
        if (Tmp1) { free (Tmp1); Tmp1 = NULL; }
        if (Tmp2) { free (Tmp2); Tmp2 = NULL; }
35 if (Tmp3) { free (Tmp3); Tmp3 = NULL; }
        if (Perm) { free (Perm); Perm = NULL; }
    }
    /*}}}*/

40 /* return the number of factors */
    static int
    factorize (int nPass, int * kt)
    {
        int nFactor = 0;
        int j, jj;
45
        *kt = 0;
        /* determine the factors of n */
        while ((nPass % 16) == 0) /* factors of 4 */
50 {
            factor [nFactor++] = 4;
            nPass /= 16;
        }
        j = 3; jj = 9; /* factors of 3, 5, 7, ... */
55 do {
            while ((nPass % jj) == 0)
            {
                factor [nFactor++] = j;
                nPass /= jj;
            }
        } while (j <= nPass);
    }

```



```

        }
        j += 2;
        jj = j * j;
    } while (jj <= nPass);
5   if (nPass <= 4)
    {
        *kt = nFactor;
        factor [nFactor] = nPass;
        if (nPass != 1)
10         nFactor++;
    }
    else
    {
        if (nPass - (nPass / 4 << 2) == 0)
15         {
            factor [nFactor++] = 2;
            nPass /= 4;
        }
        *kt = nFactor;
        j = 2;
20        do {
            if ((nPass % j) == 0)
            {
                factor [nFactor++] = j;
25                nPass /= j;
            }
            j = ((j + 1) / 2 << 1) + 1;
        } while (j <= nPass);
    }
30    if (*kt)
    {
        j = *kt;
        do
            factor [nFactor++] = factor [--j];
35        while (j);
    }

    return nFactor;
}
40

/* re-include this source file on the second pass through */
/* {{{ defines for re-including double precision */
#ifndef FFT_NODOUBLE
5   #ifndef FFT_FLOAT
    # define FFT_FLOAT
45  # endif
    #else
    # undef REAL
    # undef FFTN
50  # undef FFTNS
    # undef FFTRADIX
    # undef FFTRADIXS
    /* defines for double */
    # define REAL          double
55  # define FFTN           fftn
    # define FFTNS         "fftn"
    # define FFTRADIX      fftradix
    # define FFTRADIXS     "fftradx"
    /* double precision routine */

```

```

static int
fftradix (double Re[], double Im[],
          size_t nTotal, size_t nPass, size_t nSpan, int isign,
          int maxFactors, int maxPerm);
5  # include __FILE__          /* include this file again */
   #endif
   /*}}}*/

   /*{{{ defines for re-including float precision */
10  #ifdef FFT_FLOAT
   # undef REAL
   # undef FFTN
   # undef FFTNS
   # undef FFTRADIX
15  # undef FFTRADIXS
   /* defines for float */
   # define REAL          float
   # define FFTN          fftnf          /* trailing 'f' for float */
   # define FFTNS          "fftnf"      /* name for error message */
20  # define FFTRADIX      fftradixf /* trailing 'f' for float */
   # define FFTRADIXS      "fftradixf"  /* name for error message */
   /* float precision routine */
   static int
   fftradixf (float Re[], float Im[],
25         size_t nTotal, size_t nPass, size_t nSpan, int isign,
          int maxFactors, int maxPerm);
   # include __FILE__          /* include this file again */
   #endif
   /*}}}*/
30  #else /* _FFTN_C */

   /*
   * use macros to access the Real/Imaginary parts so that it's possible
   * to substitute different macros if a complex struct is used
35  */

   #ifndef Re_Data
   # define Re_Data(i)      Re[i]
   # define Im_Data(i)      Im[i]
40  #endif

   /*
   *
   */
45  int
   FFTN (int ndim,
        const int dims [],
        REAL Re [],
        REAL Im [],
50  int iSign,
        double scaling)
   {
       size_t nTotal;
       int maxFactors, maxPerm;
55

       /*
       * tally the number of elements in the data array
       * and determine the number of dimensions
       */

```

```
nTotal = 1;
if (ndim)
{
    if (dims != NULL)
    5      {
        int i;
        /* number of dimensions was specified */
        for (i = 0; i < ndim; i++)
        {
            10      if (dims [i] <= 0) goto Dimension_Error;
                nTotal *= dims [i];
        }
    }
    else
    15      nTotal *= ndim;
}
else
{
    int i;
    20      /* determine # of dimensions from zero-terminated list */
    if (dims == NULL) goto Dimension_Error;
    for (ndim = i = 0; dims [i]; i++)
    {
        if (dims [i] <= 0)
        25      goto Dimension_Error;
        nTotal *= dims [i];
        ndim++;
    }
}
30
/* determine maximum number of factors and permutations */
#ifdef 1
/*
    35      * follow John Beale's example, just use the largest dimension and don't
    * worry about excess allocation. May be someone else will do it?
    */
    if (dims != NULL)
    {
        int i;
        40      for (maxFactors = maxPerm = 1, i = 0; i < ndim; i++)
        {
            if (dims [i] > maxFactors) maxFactors = dims [i];
            if (dims [i] > maxPerm) maxPerm = dims [i];
        }
    }
    45      else
    {
        maxFactors = maxPerm = nTotal;
    }
#endif
50
/* use the constants used in the original Fortran code */
maxFactors = 23;
maxPerm = 209;
#endif
55
/* loop over the dimensions: */
if (dims != NULL)
{
    size_t nSpan = 1;
    int i;
```

```

        for (i = 0; i < ndim; i++)
        {
            int ret;
5           nSpan *= dims [i];
            ret = FFTRADIX (Re, Im, nTotal, dims [i], nSpan, iSign,
                           maxFactors, maxPerm);
            /* exit, clean-up already done */
            if (ret)
10             return ret;
        }
    }
    else
    {
15         int ret;
        ret = FFTRADIX (Re, Im, nTotal, nTotal, nTotal, iSign,
                        maxFactors, maxPerm);
        /* exit, clean-up already done */
        if (ret)
20         return ret;
    }

    /* Divide through by the normalizing constant: */
    if (scaling && scaling != 1.0)
25     {
        int i;

        if (iSign < 0) iSign = -iSign;
        if (scaling < 0.0)
30         scaling = (scaling < -1.0) ? sqrt (nTotal) : nTotal;
        scaling = 1.0 / scaling;    /* multiply is often faster */
        for (i = 0; i < nTotal; i += iSign)
        {
            Re_Data (i) *= scaling;
35             Im_Data (i) *= scaling;
        }
    }
    return 0;

40     Dimension_Error:
    fprintf (stderr, "Error: " FFTNS "(" - dimension error\n");
    fft_free ();    /* free-up memory */
    return -1;
}

45  /*-----*/
/*
 * singleton's mixed radix routine
 *
50  * could move allocation out to fftn(), but leave it here so that it's
 * possible to make this a standalone function
 */
static int
FFTRADIX (REAL Re [],
55         REAL Im [],
         size_t nTotal,
         size_t nPass,
         size_t nSpan,
         int iSign,

```

```

        int maxFactors,
        int maxPerm)
    {
        int ii, nFactor, kspan, ispan, inc;
5      int j, jc, jf, jj, k, k1, k3, kk, kt, nn, ns, nt;

        REAL radf;
        REAL c1, c2, c3, cd;
        REAL s1, s2, s3, sd;

10     REAL * Rtmp = NULL;      /* temp space for real part */
        REAL * Itmp = NULL;    /* temp space for imaginary part */
        REAL * Cos = NULL;     /* Cosine values */
        REAL * Sin = NULL;     /* Sine values */

15     #ifndef FFT_RADIX4
        REAL s60 = SIN60;      /* sin(60 deg) */
        REAL s72 = SIN72;      /* sin(72 deg) */
        REAL c72 = COS72;      /* cos(72 deg) */
20     #endif
        REAL pi2 = M_PI;       /* use PI first, 2 PI later */

        /* gcc complains about k3 being uninitialized, but I can't find out where
        * or why ... it looks okay to me.
25     *
        * initialize to make gcc happy
        */
        k3 = 0;

30     /* gcc complains about c2, c3, s2,s3 being uninitialized, but they're
        * only used for the radix 4 case and only AFTER the (s1 == 0.0) pass
        * through the loop at which point they will have been calculated.
        *
        * initialize to make gcc happy
35     */
        c2 = c3 = s2 = s3 = 0.0;

        /* Parameter adjustments, was fortran so fix zero-offset */
        Re--;
40     Im--;

        if (nPass < 2)
            return 0;

45     /* allocate storage */
        if (SpaceAlloced < maxFactors * sizeof (REAL))
        {
            #ifndef SUN_BROKEN_REALLOC
            if (!SpaceAlloced) /* first time */
50             {
                SpaceAlloced = maxFactors * sizeof (REAL);
                Tmp0 = malloc (SpaceAlloced);
                Tmp1 = malloc (SpaceAlloced);
                Tmp2 = malloc (SpaceAlloced);
55             Tmp3 = malloc (SpaceAlloced);
            }
            else
            {
                #endif
    #endif

```

```

        SpaceAlloced = maxFactors * sizeof (REAL);
        Tmp0 = realloc (Tmp0, SpaceAlloced);
        Tmp1 = realloc (Tmp1, SpaceAlloced);
        Tmp2 = realloc (Tmp2, SpaceAlloced);
        Tmp3 = realloc (Tmp3, SpaceAlloced);
5      #ifdef SUN_BROKEN_REALLOC
        }
      #endif
10     else
        {
            /* allow full use of alloc'd space */
            maxFactors = SpaceAlloced / sizeof (REAL);
        }
15     if (MaxPermAlloced < maxPerm)
        {
            #ifdef SUN_BROKEN_REALLOC
                if (!MaxPermAlloced) /* first time */
                    Perm = malloc (maxPerm * sizeof(int));
20                else
            #endif
                Perm = realloc (Perm, maxPerm * sizeof(int));
                MaxPermAlloced = maxPerm;
            }
25     else
        {
            /* allow full use of alloc'd space */
            maxPerm = MaxPermAlloced;
        }
30     if (!Tmp0 || !Tmp1 || !Tmp2 || !Tmp3 || !Perm) goto Memory_Error;

        /* assign pointers */
        Rtmp = (REAL *) Tmp0;
        Itmp = (REAL *) Tmp1;
35     Cos = (REAL *) Tmp2;
        Sin = (REAL *) Tmp3;

        /*
        * Function Body
        */
40     inc = iSign;
        if (iSign < 0)
        {
            #ifndef FFT_RADIX4
45             s60 = -s60;
             s72 = -s72;
            #endif
            pi2 = -pi2;
            inc = -inc; /* absolute value */
50     }

        /* adjust for strange increments */
        nt = inc * nTotal;
        ns = inc * nSpan;
55     kspan = ns;

        nn = nt - inc;
        jc = ns / nPass;
        radf = pi2 * (double) jc;
    
```

```

pi2 *= 2.0;                                /* use 2 PI from here on */

ii = 0;
jf = 0;
5  /* determine the factors of n */

nFactor = factorize (nPass, &kt);
/* test that nFactors is in range */
if (nFactor > NFACTOR)
10  {
    fprintf(stderr, "Error: " FFTRADIXS "() - exceeded number of factors\n");
    goto Memory_Error;
}

15  /* compute fourier transform */
for (;;) {
    sd = radf / (double) kspan;
    cd = sin (sd);
    cd = 2.0 * cd * cd;
20  sd = sin (sd + sd);
    kk = 1;
    ii++;

    switch (factor [ii - 1]) {
25  case 2:
        /* transform for factor of 2 (including rotation factor) */
        kspan /= 2;
        k1 = kspan + 2;
        do {
30          do {
              REAL tmp;
              REAL tmpi;
              int k2;

35          k2 = kk + kspan;
              tmp = Re_Data (k2);
              tmpi = Im_Data (k2);
              Re_Data (k2) = Re_Data (kk) - tmp;
              Im_Data (k2) = Im_Data (kk) - tmpi;
40          Re_Data (kk) += tmp;
              Im_Data (kk) += tmpi;
              kk = k2 + kspan;
          } while (kk <= nn);
          kk -= nn;
45        } while (kk <= jc);
        if (kk > kspan)
            goto Permute_Results; /* exit infinite loop */
        do {
50          int k2;

          c1 = 1.0 - cd;
          s1 = sd;
          do {
55          REAL tmp;
              do {
                  do {
                      REAL tmp;
                      REAL tmpi;

```

```

        k2 = kk + kspan;
        tmpr = Re_Data(kk) - Re_Data(k2);
        tmpi = Im_Data(kk) - Im_Data(k2);
        Re_Data(kk) += Re_Data(k2);
5       Im_Data(kk) += Im_Data(k2);
        Re_Data(k2) = c1 * tmpr - s1 * tmpi;
        Im_Data(k2) = s1 * tmpr + c1 * tmpi;
        kk = k2 + kspan;
    } while (kk < nt);
10     k2 = kk - nt;
        c1 = -c1;
        kk = k1 - k2;
    } while (kk > k2);
        tmp = c1 - (cd * c1 + sd * s1);
15     s1 = sd * c1 - cd * s1 + s1;
        c1 = 2.0 - (tmp * tmp + s1 * s1);
        s1 *= c1;
        c1 *= tmp;
        kk += jc;
20     } while (kk < k2);
        k1 += (inc + inc);
        kk = (k1 - kspan) / 2 + jc;
    } while (kk <= jc + jc);
    break;
25
case 4:                                     /* transform for factor of 4 */
    ispan = kspan;
    kspan /= 4;
30     do {
        c1 = 1.0;
        s1 = 0.0;
        do {
            do {
35                 REAL ajm, ajp, akm, akp;
                    REAL bjm, bjp, bkm, bkp;
                    int k2;

                    k1 = kk + kspan;
40                 k2 = k1 + kspan;
                    k3 = k2 + kspan;
                    akp = Re_Data(kk) + Re_Data(k2);
                    akm = Re_Data(kk) - Re_Data(k2);

45                 ajp = Re_Data(k1) + Re_Data(k3);
                    ajm = Re_Data(k1) - Re_Data(k3);

                    bkp = Im_Data(kk) + Im_Data(k2);
                    bkm = Im_Data(kk) - Im_Data(k2);
50                 bjp = Im_Data(k1) + Im_Data(k3);
                    bjm = Im_Data(k1) - Im_Data(k3);

                    Re_Data(kk) = akp + ajp;
55                 Im_Data(kk) = bkp + bjp;
                    ajp = akp - ajp;
                    bjp = bkp - bjp;
                    if (iSign < 0)
                    {

```



```

        akp = akm + bjm;
        bkp = bkm - ajm;
        akm -= bjm;
        bkm += ajm;
5      }
      else
      {
        akp = akm - bjm;
        bkp = bkm + ajm;
10     akm += bjm;
        bkm -= ajm;
      }
      /* avoid useless multiplies */
      if (s1 == 0.0)
15     {
        Re_Data (k1) = akp;
        Re_Data (k2) = ajp;
        Re_Data (k3) = akm;
        Im_Data (k1) = bkp;
20     Im_Data (k2) = bjp;
        Im_Data (k3) = bkm;
      }
      else
      {
25     Re_Data (k1) = akp * c1 - bkp * s1;
        Re_Data (k2) = ajp * c2 - bjp * s2;
        Re_Data (k3) = akm * c3 - bkm * s3;
        Im_Data (k1) = akp * s1 + bkp * c1;
        Im_Data (k2) = ajp * s2 + bjp * c2;
30     Im_Data (k3) = akm * s3 + bkm * c3;
      }
      kk = k3 + kspan;
    } while (kk <= nt);

35   c2 = c1 - (cd * c1 + sd * s1);
      s1 = sd * c1 - cd * s1 + s1;
      c1 = 2.0 - (c2 * c2 + s1 * s1);
      s1 *= c1;
      c1 *= c2;
40   /* values of c2, c3, s2, s3 that will get used next time */
      c2 = c1 * c1 - s1 * s1;
      s2 = 2.0 * c1 * s1;
      c3 = c2 * c1 - s2 * s1;
      s3 = c2 * s1 + s2 * c1;
45   kk = kk - nt + jc;
    } while (kk <= kspan);
      kk = kk - kspan + inc;
    } while (kk <= jc);
    if (kspan == jc)
50   goto Permute_Results; /* exit infinite loop */
    break;

default:
    /* transform for odd factors */
55   #ifdef FFT_RADIX4
      fprintf (stderr, "Error: " FFTRADIXS "(): compiled for radix 2/4 only\n");
      fft_free (); /* free-up memory */
      return -1;
      break;

```

```

5  #else /* FFT_RADIX4 */
    ispan = kspan;
    k = factor [ii - 1];
    kspan /= factor [ii - 1];

    switch (factor [ii - 1]) {
    case 3: /* transform for factor of 3 (optional code) */
        do {
            do {
10         REAL aj, tmpr;
            REAL bj, tmpi;
            int k2;

            k1 = kk + kspan;
15         k2 = k1 + kspan;
            tmpr = Re_Data (kk);
            tmpi = Im_Data (kk);
            aj = Re_Data (k1) + Re_Data (k2);
            bj = Im_Data (k1) + Im_Data (k2);
20         Re_Data (kk) = tmpr + aj;
            Im_Data (kk) = tmpi + bj;
            tmpr -= 0.5 * aj;
            tmpi -= 0.5 * bj;
            aj = (Re_Data (k1) - Re_Data (k2)) * s60;
25         bj = (Im_Data (k1) - Im_Data (k2)) * s60;
            Re_Data (k1) = tmpr - bj;
            Re_Data (k2) = tmpr + bj;
            Im_Data (k1) = tmpi + aj;
            Im_Data (k2) = tmpi - aj;
30         kk = k2 + kspan;
            } while (kk < nn);
            kk -= nn;
        } while (kk <= kspan);
        break;

35     case 5: /* transform for factor of 5 (optional code) */
        c2 = c72 * c72 - s72 * s72;
        s2 = 2.0 * c72 * s72;
        do {
40         do {
            REAL aa, aj, ak, ajm, ajp, akm, akp;
            REAL bb, bj, bk, bjmi, bjp, bkm, bkp;
            int k2, k4;

45         k1 = kk + kspan;
            k2 = k1 + kspan;
            k3 = k2 + kspan;
            k4 = k3 + kspan;
            akp = Re_Data (k1) + Re_Data (k4);
50         akm = Re_Data (k1) - Re_Data (k4);
            bkp = Im_Data (k1) + Im_Data (k4);
            bkm = Im_Data (k1) - Im_Data (k4);
            ajp = Re_Data (k2) + Re_Data (k3);
            ajm = Re_Data (k2) - Re_Data (k3);
55         bjp = Im_Data (k2) + Im_Data (k3);
            bjmi = Im_Data (k2) - Im_Data (k3);
            aa = Re_Data (kk);
            bb = Im_Data (kk);
            Re_Data (kk) = aa + akp + ajp;
    
```

```

        Im_Data (kk) = bb + bkp + bjp;
        ak = akp * c72 + ajp * c2 + aa;
        bk = bkp * c72 + bjp * c2 + bb;
        aj = akm * s72 + ajm * s2;
5      bj = bkm * s72 + bjm * s2;
        Re_Data (k1) = ak - bj;
        Re_Data (k4) = ak + bj;
        Im_Data (k1) = bk + aj;
        Im_Data (k4) = bk - aj;
10     ak = akp * c2 + ajp * c72 + aa;
        bk = bkp * c2 + bjp * c72 + bb;
        aj = akm * s2 - ajm * s72;
        bj = bkm * s2 - bjm * s72;
15     Re_Data (k2) = ak - bj;
        Re_Data (k3) = ak + bj;
        Im_Data (k2) = bk + aj;
        Im_Data (k3) = bk - aj;
        kk = k4 + kspan;
    } while (kk < nn);
20     kk -= nn;
    } while (kk <= kspan);
    break;

default:
25     k = factor [ii - 1];
    if (jf != k)
    {
        jf = k;
        s1 = pi2 / (double) jf;
30     c1 = cos (s1);
        s1 = sin (s1);
        if (jf > maxFactors)
            goto Memory_Error;
        Cos [jf - 1] = 1.0;
35     Sin [jf - 1] = 0.0;
        j = 1;
        do {
            Cos [j - 1] = Cos [k - 1] * c1 + Sin [k - 1] * s1;
            Sin [j - 1] = Cos [k - 1] * s1 - Sin [k - 1] * c1;
40     k--;
            Cos [k - 1] = Cos [j - 1];
            Sin [k - 1] = -Sin [j - 1];
            j++;
        } while (j < k);
45     }
    do {
        do {
            REAL aa, ak;
            REAL bb, bk;
50     int k2;

            aa = ak = Re_Data (kk);
            bb = bk = Im_Data (kk);

55     k1 = kk;
            k2 = kk + ispan;
            j = 1;
            k1 += kspan;
            do {

```

```

        k2 := kspan;
        Rtmp [j] = Re_Data (k1) + Re_Data (k2);
        ak += Rtmp [j];
        Itmp [j] = Im_Data (k1) + Im_Data (k2);
5         bk += Itmp [j];
        j++;
        Rtmp [j] = Re_Data (k1) - Re_Data (k2);
        Itmp [j] = Im_Data (k1) - Im_Data (k2);
        j++;
10         k1 += kspan;
    } while (k1 < k2);
    Re_Data (kk) = ak;
    Im_Data (kk) = bk;

15         k1 = kk;
        k2 = kk + ispan;
        j = 1;
        do {
20             REAL aj = 0.0;
             REAL bj = 0.0;

            k1 += kspan;
            k2 := kspan;
            jj = j;
25             ak = aa;
            bk = bb;
            k = 1;
            do {
30                 ak += Rtmp [k] * Cos [jj - 1];
                 bk += Itmp [k] * Cos [jj - 1];
                 k++;
                 aj += Rtmp [k] * Sin [jj - 1];
                 bj += Itmp [k] * Sin [jj - 1];
35                 k++;
                 jj += j;
                 if (jj > jf)
                     jj -= jf;
            } while (k < jf);
            k = jf - j;
40             Re_Data (k1) = ak - bj;
            Im_Data (k1) = bk + aj;
            Re_Data (k2) = ak + bj;
            Im_Data (k2) = bk - aj;
            j++;
45         } while (j < k);
        kk += ispan;
    } while (kk <= nn);
    kk := nn;
    } while (kk <= kspan);
50     break;
}
/* multiply by rotation factor (except for factors of 2 and 4) */
if (ii == nFactor)
    goto Permute_Results; /* exit infinite loop */
55     kk = jc + 1;
    do {
        c2 = 1.0 - cd;
        s1 = sd;
        do {

```

```

        c1 = c2;
        s2 = s1;
        kk += kspan;
        do {
5           REAL tmp;
            do {
                REAL ak;
                ak = Re_Data (kk);
                Re_Data (kk) = c2 * ak - s2 * Im_Data (kk);
10             Im_Data (kk) = s2 * ak + c2 * Im_Data (kk);
                kk += ispan;
            } while (kk <= nt);
            tmp = s1 * s2;
            s2 = s1 * c2 + c1 * s2;
15             c2 = c1 * c2 - tmp;
            kk = kk - nt + kspan;
        } while (kk <= ispan);
        c2 = c1 - (cd * c1 + sd * s1);
        s1 += sd * c1 - cd * s1;
20         c1 = 2.0 - (c2 * c2 + s1 * s1);
        s1 *= c1;
        c2 *= c1;
        kk = kk - ispan + jc;
    } while (kk <= kspan);
25     kk = kk - kspan + jc + inc;
    } while (kk <= jc + jc);
    break;
#endif /* FFT_RADIX4 */
    }
30 }

/* permute the results to normal order -- done in two stages */
/* permutation for square factors of n */
Permute_Results:
35     Perm[0] = ns;
    if (kt)
    {
        int k2;

40         k = kt + kt + 1;
        if (k > nFactor)
            k--;
        Perm[k] = jc;
        j = 1;
45         do {
            Perm[j] = Perm[j - 1] / factor[j - 1];
            Perm[k - 1] = Perm[k] * factor[j - 1];
            j++;
            k--;
50         } while (j < k);
        k3 = Perm[k];
        kspan = Perm[1];
        kk = jc + 1;
        k2 = kspan + 1;
55         j = 1;
        if (nPass != nTotal)
        {
            /* permutation for multivariate transform */
            Permute_Multi:

```

```

do {
    do {
        k = kk + jc;
        do {
            /* swap
5          * Re_Data (kk) <> Re_Data (k2)
            * Im_Data (kk) <> Im_Data (k2)
            */
            REAL tmp;
10         tmp = Re_Data (kk); Re_Data (kk) = Re_Data (k2); Re_Data (k2) = tmp;
            tmp = Im_Data (kk); Im_Data (kk) = Im_Data (k2); Im_Data (k2) = tmp;
            kk += inc;
            k2 += inc;
        } while (kk < k);
15        kk += (ns - jc);
        k2 += (ns - jc);
    } while (kk < nt);
    k2 = k2 - nt + kspan;
    kk = kk - nt + jc;
20 } while (k2 < ns);
do {
    do {
        k2 -= Perm [j - 1];
        j++;
25        k2 = Perm [j] + k2;
    } while (k2 > Perm [j - 1]);
    j = 1;
    do {
        if (kk < k2)
30        goto Permute_Multi;
        kk += jc;
        k2 += kspan;
    } while (k2 < ns);
} while (kk < ns);
35 }
else
{
    /* permutation for single-variate transform (optional code) */
    Permute_Single:
40    do {
        /* swap
        * Re_Data (kk) <> Re_Data (k2)
        * Im_Data (kk) <> Im_Data (k2)
        */
        REAL t;
45        t = Re_Data (kk); Re_Data (kk) = Re_Data (k2); Re_Data (k2) = t;
        t = Im_Data (kk); Im_Data (kk) = Im_Data (k2); Im_Data (k2) = t;
        kk += inc;
        k2 += kspan;
50    } while (k2 < ns);
    do {
        do {
            k2 -= Perm [j - 1];
            j++;
55            k2 = Perm [j] + k2;
        } while (k2 > Perm [j - 1]);
        j = 1;
        do {
            if (kk < k2)

```

```

        goto Permute_Single;
        kk += inc;
        k2 += kspan;
    } while (k2 < ns);
5    } while (kk < ns);
    }
    jc = k3;
}

10  if ((kt << 1) + 1 >= nFactor)
    return 0;
    ispan = Perm [kt];

    /* permutation for square-free factors of n */
15  j = nFactor - kt;
    factor [j] = 1;
    do {
        factor [j - 1] *= factor [j];
        j--;
20  } while (j != kt);
    nn = factor [kt] - 1;
    kt++;
    if (nn > maxPerm)
        goto Memory_Error;
25  j = jj = 0;
    for (;;) {
        int k2;

30  k = kt + 1;
        k2 = factor [kt - 1];
        kk = factor [k - 1];
        j++;
        if (j > nn)
35  break;
        jj += kk;
        while (jj >= k2)
        {
            jj -= k2;
            k2 = kk;
            kk = factor [k++];
            jj += kk;
40  }
        Perm [j - 1] = jj;
45  }
    /* determine the permutation cycles of length greater than 1 */
    j = 0;
    for (;;) {
        do {
50  kk = Perm [j++];
        } while (kk < 0);
        if (kk != j)
        {
85  do {
            k = kk;
            kk = Perm [k - 1];
            Perm [k - 1] = -kk;
        } while (kk != j);
            k3 = kk;

```

```

    }
    else
    {
        Perm[j - 1] = -j;
5      if (j == nn)
        break;          /* exit infinite loop */
    }
}

10  maxFactors *= inc;

/* reorder a and b, following the permutation cycles */
for (;;) {
    j = k3 + 1;
15  nt := ispan;
    ii = nt - inc + 1;
    if (nt < 0)
        break;          /* exit infinite loop */
    do {
20      do {
          j--;
        } while (Perm[j - 1] < 0);
        jj = jc;
        do {
25          int k2;

          if (jj < maxFactors) kspan = jj; else kspan = maxFactors;

          jj -= kspan;
30          k = Perm[j - 1];
          kk = jc * k + ii + jj;

          k1 = kk + kspan;
          k2 = 0;
35          do {
              Rtmp[k2] = Re_Data(k1);
              Itmp[k2] = Im_Data(k1);
              k2++;
              k1 -= inc;
40          } while (k1 != kk);

          do {
              k1 = kk + kspan;
              k2 = k1 - jc * (k + Perm[k - 1]);
45          k = -Perm[k - 1];
              do {
                  Re_Data(k1) = Re_Data(k2);
                  Im_Data(k1) = Im_Data(k2);
                  k1 -= inc;
50                  k2 -= inc;
              } while (k1 != kk);
              kk = k2;
          } while (k != jj);

55          k1 = kk + kspan;
          k2 = 0;
          do {
              Re_Data(k1) = Rtmp[k2];
              Im_Data(k1) = Itmp[k2];

```



```
        k2++;
        k1 -= inc;
        } while (k1 != kk);
    } while (jj);
5    } while (j != 1);
    }
    return 0;                      /* exit point here */

    /* alloc or other problem, do some clean-up */
10 Memory_Error:
    fprintf(stderr, "Error: " FFTRADIXS "() - insufficient memory.\n");
    fft_free ();                  /* free-up memory */
    return -1;
}
15 #endif /* _FFTN_C */
/*----- end-of-file (C source) -----*/
```

```

/*-----*C*-----*
* File:
*      fftn.h
*
5  * Singleton's multivariate complex Fourier transform, computed in
* place using mixed-radix Fast Fourier Transform algorithm.
*
* Called here 'fftn' since it does a radix-n FFT on n-dimensional data
*
10 * Copyright(c)1995,97 Mark Olesen <olesen@mc.QueensU.CA>
*      Queen's Univ at Kingston (Canada)
*
* Permission to use, copy, modify, and distribute this software for
* any purpose without fee is hereby granted, provided that this
15 * entire notice is included in all copies of any software which is
* or includes a copy or modification of this software and in all
* copies of the supporting documentation for such software.
*
* THIS SOFTWARE IS BEING PROVIDED "AS IS", WITHOUT ANY EXPRESS OR
20 * IMPLIED WARRANTY. IN PARTICULAR, NEITHER THE AUTHOR NOR QUEEN'S
* UNIVERSITY AT KINGSTON MAKES ANY REPRESENTATION OR WARRANTY OF ANY
* KIND CONCERNING THE MERCHANTABILITY OF THIS SOFTWARE OR ITS
* FITNESS FOR ANY PARTICULAR PURPOSE.
*
25 * All of which is to say that you can do what you like with this
* source code provided you don't try to sell it as your own and you
* include an unaltered copy of this message (including the
* copyright).
*
30 * It is also implicitly understood that bug fixes and improvements
* should make their way back to the general Internet community so
* that everyone benefits.
*
* Brief overview of parameters:
35 * -----*
* Re[]: real value array
* Im[]: imaginary value array
* nTotal:      total number of complex values
* nPass: number of elements involved in this pass of transform
40 * nSpan:      nspan/nPass = number of bytes to increment pointer
*              in Re[] and Im[]
* isign: exponent: +1 = forward -1 = reverse
* scaling:      normalizing constant by which the final result is DIVIDED
* scaling == -1, normalize by total dimension of the transform
45 * scaling < -1, normalize by the square-root of the total dimension
*
* Slightly more detailed information:
* -----*
50 * void fft_free (void);
*
* free-up allocated temporary storage after finished all the Fourier
* transforms.
*
55 * -----*
* int fftn (int ndim, const int dims[], REAL Re[], REAL Im[],
*          int iSign, double scaling);
*

```

```

* NDIM = the total number dimensions
* DIMS = a vector of array sizes
*     if NDIM is zero then DIMS must be zero-terminated
*
5  * RE and IM hold the real and imaginary components of the data, and
*   return the resulting real and imaginary Fourier coefficients.
* Multidimensional data *must* be allocated contiguously. There is
* no limit on the number of dimensions.
*
10 * ISIGN = the sign of the complex exponential
*     (ie, forward or inverse FFT)
*     the magnitude of ISIGN (normally 1) is used to determine
*     the correct indexing increment (see below).
*
15 * SCALING = normalizing constant by which the final result is DIVIDED
*     if SCALING == -1, normalize by total dimension of the transform
*     if SCALING < -1, normalize by the square-root of the total dimension
*
* example:
20 * tri-variate transform with Re[n3][n2][n1], Im[n3][n2][n1]
*
*     int dims[3] = {n1,n2,n3}
*     fftn (3, dims, Re, Im, 1, scaling);
*
25 * or, using a null terminated dimension list
*     int dims[4] = {n1,n2,n3,0}
*     fftn (0, dims, Re, Im, 1, scaling);
* -----*/
#ifndef _FFTN_H
30 #define _FFTN_H
#ifdef __cplusplus
extern "C" {
#endif
35     extern void fft_free (void);

    /* double precision routine */
    extern int fftn (int /* ndim */,
                    const int /* dims */[],
                    double /* Re */[],
40                    double /* Im */[],
                    int /* isign */,
                    double /* scaling */);

    /* float precision routine */
55     extern int fftnf (int /* ndim */,
                      const int /* dims */[],
                      float /* Re */[],
                      float /* Im */[],
                      int /* isign */,
                      double /* scaling */);

#ifdef __cplusplus
}
#endif
#endif /* _FFTN_H */
/*----- end-of-file (C header) -----*/

```